VISIT…

ББК 22.19
С17
УДК 519.6

РФФИ

Издание осуществлено при поддержке Российского фонда фундаментальных исследований по проекту 95-01-21063

Самарский А.А., Михайлов А.П. Математическое моделирование: Идеи. Методы. Примеры. — М.: Наука. Физматлит, 1997. — 320 с. — ISBN 5-02-015186-6.

В монографии изложены универсальные методологические подходы, позволяющие безотносительно к конкретным областям приложений строить адекватные математические модели изучаемых объектов. Представлены методы и примеры построения и анализа математических моделей для различных задач механики, физики, биологии, экономики, социологии на основе использования фундаментальных законов природы, вариационных принципов, иерархических цепочек, метода аналогий.

Для специалистов по математическому моделированию, аспирантов, студентов, изучающих и использующих методы математического моделирования, вычислительного эксперимента.

Табл. 5. Ил. 123. Библиогр. 89 назв.

Научное издание

САМАРСКИЙ Александр Андреевич,
МИХАЙЛОВ Александр Петрович

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ: ИДЕИ. МЕТОДЫ. ПРИМЕРЫ

Редактор *Е.Ю.Ходан*
Компьютерный набор *А.С.Болдарева, Н.Г.Сиротенко*

ИБ № 41787

ЛР № 020297 от 27.11.91. Подписано в печать с оригинал-макета 25.06.97.
Формат 60 × 90/16. Бумага офсетная № 1. Печать офсетная.
Усл. печ. л. 20. Уч.-изд. л. 22. Тираж 2000 экз. Заказ тип. № 1922 . С–018.

Издательская фирма «Физико-математическая литература» РАН
117071 Москва В-71, Ленинский проспект, 15

Отпечатано в Московской типографии № 2 РАН
121099 Москва Г-99, Шубинский пер., 6

С $\frac{1402020000-018}{053(02)-97}$ 62-97. Наука, II полугодие

© А.А.Самарский, А.П.Михайлов, 1997

ISBN 5-02-015186-6

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ 6

ГЛАВА I

ПРОСТЕЙШИЕ МАТЕМАТИЧЕСКИЕ МОДЕЛИ И ОСНОВНЫЕ ПОНЯТИЯ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ

§ 1. Элементарные математические модели 11

1. Фундаментальные законы природы (11). 2. Вариационные принципы (16). 3. Применение аналогий при построении моделей (19) 4. Иерархический подход к получению моделей (21). 5. О нелинейности математических моделей (23). 6. Предварительные выводы (25). Упражнения (26).

§ 2. Примеры моделей, получаемых из фундаментальных законов природы 26

1. Траектория всплытия подводной лодки (26). 2. Отклонение заряженной частицы в электронно-лучевой трубке (28). 3. Колебания конец Сатурна (30). 4. Движение шарика, присоединенного к пружине (32). 5. Заключение (33). Упражнения (34).

§ 3. Вариационные принципы и математические модели 34

1. Общая схема принципа Гамильтона (34). 2. Третий способ получения модели системы «шарик—пружина» (35). 3. Колебания маятника в поле сил тяжести (37). 4. Заключение (39). Упражнения (40).

§ 4. Пример иерархии моделей 40

1. Различные варианты действия заданной внешней силы (40). 2. Движение точки крепления, пружина на вращающемся стержне (41). 3. Учет сил трения (42). 4. Два типа нелинейных моделей системы «шарик—пружина» (44). 5. Заключение (47). Упражнения (47)

§ 5. Универсальность математических моделей 48

1. Жидкость в U-образном сосуде (48). 2. Колебательный электрический контур (49). 3 Малые колебания при взаимодействии двух биологических популяций (50). 4 Простейшая модель изменения зарплаты и занятости (51). 5. Заключение (52). Упражнения (52).

§ 6. Некоторые модели простейших нелинейных объектов 53

1. О происхождении нелинейности (53). 2. Три режима в нелинейной модели популяции (53). 3. Влияние сильной нелинейности на процесс колебаний (55). 4. О численных методах (56). Упражнения (57).

ГЛАВА II

ПОЛУЧЕНИЕ МОДЕЛЕЙ ИЗ ФУНДАМЕНТАЛЬНЫХ ЗАКОНОВ ПРИРОДЫ

§ 1. Сохранение массы вещества 58

1 Поток частиц в трубе (58). 2. Основные предположения о гравитационном режиме течения грунтовых вод (60). 3 Баланс массы в элементе грунта (61). 4. Замыкание закона сохранения массы (63). 5 О некоторых свойствах уравнения Буссинеска (64) Упражнения (66).

§ 2. Сохранение энергии 66

1. Предварительные сведения о процессах теплопередачи (66). 2. Вывод закона Фурье из молекулярно-кинетических представлений (68). 3. Уравнение баланса тепла (70) 4. Постановка типичных краевых условий для уравнения теплопроводности (72) 5 Об особенностях моделей теплопередачи (74). Упражнения (76).

§ 3. Сохранение числа частиц 76

1. Основные понятия теории теплового излучения (76). 2. Уравнение баланса числа фотонов в среде (78). 3 Некоторые свойства уравнения переноса излучения (80). Упражнения (82).

§ 4. Совместное применение нескольких фундаментальных законов . . . 82
1 Предварительные понятия газовой динамики (82) 2 Уравнение неразрывности для сжимаемого газа (83). 3 Уравнения движения газа (84) 4. Уравнение энергии (85) 5 Уравнения газовой динамики в лагранжевых координатах (86). 6 Краевые условия для уравнений газовой динамики (88). 7 Некоторые особенности моделей газовой динамики (90). Упражнения (92).

ГЛАВА III
МОДЕЛИ ИЗ ВАРИАЦИОННЫХ ПРИНЦИПОВ, ИЕРАРХИИ МОДЕЛЕЙ

§ 1. Уравнения движения, вариационные принципы и законы сохранения в механике . . . 93
1. Уравнения движения механической системы в форме Ньютона (93). 2. Уравнения движения в форме Лагранжа (95). 3. Вариационный принцип Гамильтона (99). 4. Законы сохранения и свойства пространства-времени (101). Упражнения (105)
§ 2. Модели некоторых механических систем . . . 105
1. Маятник на свободной подвеске (105). 2. Непотенциальные колебания (109). 3. Малые колебания струны (111). 4. Электромеханическая аналогия (115). Упражнения (117).
§ 3. Уравнение Больцмана и производные от него . . . 118
1 Описание совокупности частиц с помощью функции распределения (118). 2 Уравнение Больцмана для функции распределения (120). 3. Распределение Максвелла и H-теорема (121) 4. Уравнения для моментов функции распределения (124). 5 Цепочка гидродинамических моделей газа (130). Упражнения (135).

ГЛАВА IV
МОДЕЛИ НЕКОТОРЫХ ТРУДНОФОРМАЛИЗУЕМЫХ ОБЪЕКТОВ

§ 1. Универсальность математических моделей . . . 136
1. Динамика скопления амеб (136). 2. Случайный марковский процесс (140). 3. Примеры аналогий между механическими, термодинамическими и экономическими объектами (146). Упражнения (150).
§ 2. Некоторые модели финансовых и экономических процессов . . . 150
1 Организация рекламной кампании (150). 2. Взаимозачет долгов предприятий (154) 3. Макромодель равновесия рыночной экономики (160). 4. Макромодель экономического роста (167). Упражнения (169).
§ 3. Некоторые модели соперничества . . . 170
1. Взаимоотношения в системе «хищник—жертва» (171). 2. Гонка вооружений между двумя странами (173). 3. Боевые действия двух армий (175). Упражнения (179).
§ 4. Динамика распределения власти в иерархии . . . 180
1 Общая постановка задачи и терминологии (180). 2. Механизмы перераспределения власти внутри иерархической структуры (186). 3. Баланс власти в инстанции, условия на ее границах и переход к непрерывной модели (189). 4. Правовая система «власть—общество». Стационарные распределения и выход власти за рамки полномочий (194). 5. Роль основных характеристик системы в феномене превышения (приниижения) власти (198). 6. Интерпретация результатов и выводы (199) Упражнения (201).

ГЛАВА V
ИССЛЕДОВАНИЕ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ

§ 1. Применение методов подобия . . . 202
1 Анализ размерностей и групповой анализ моделей (202). 2. Автомодельные (самоподобные) процессы (208) 3. Различные режимы распространения возмущений в нелинейных средах (214). Упражнения (221).
§ 2. Принцип максимума и теоремы сравнения . . . 221
1 Формулировка, некоторые следствия (221). 2. Классификация режимов с обострением (226) 3 Расширение «автомодельного метода» (229). Упражнения (234)
§ 3. Метод осреднения . . . 235
1 Локализованные структуры в нелинейных средах (235). 2. Различные способы осреднения (238) 3 Классификация режимов горения теплопроводной среды (241) Упражнения (246).

§ 4. О переходе к дискретным моделям . . . 246
1. Необходимость численного моделирования, элементарные понятия теории разностных схем (246). 2. Непосредственная формальная аппроксимация (251). 3. Интегро-интерполяционный метод (256). 4. Принцип полной консервативности (259). 5. Построение разностных схем с помощью вариационных принципов (262). 6. Использование иерархического подхода к получению дискретных моделей (265). Упражнения (269).

ГЛАВА VI
МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ СЛОЖНЫХ ОБЪЕКТОВ

§ 1. Задачи технологии и экологии . . . 270
1. Физически «безопасный» ядерный реактор (270). 2. Гидрологический «барьер» против загрязнения грунтовых вод (274). 3. Сложные режимы обтекания тел газом (277). 4. Экологически приемлемые технологии сжигания углеводородных топлив (281).
§ 2. Фундаментальные проблемы естествознания . . . 283
1. Нелинейные эффекты в лазерной термоядерной плазме (283). 2. Математическая реставрация Тунгусского феномена (288). 3. Климатические последствия ядерного конфликта (292). 4. Магнитогидродинамическое «динамо» Солнца (296).
§ 3. Вычислительный эксперимент с моделями трудноформализуемых объектов . . . 299
1 Диссипативные биологические структуры (299). 2. Процессы в переходной экономике (302). 3. Тоталитарные и анархические эволюции распределения власти в иерархиях (306).

СПИСОК ЛИТЕРАТУРЫ . . . 313

Светлой памяти академика
Андрея Николаевича Тихонова
посвящается

ВВЕДЕНИЕ

Невозможно представить себе современную науку без широкого применения математического моделирования. Сущность этой методологии состоит в замене исходного объекта его «образом» — математической моделью — и дальнейшем изучении модели с помощью реализуемых на компьютерах вычислительно-логических алгоритмов. Этот «третий метод» познания, конструирования, проектирования сочетает в себе многие достоинства как теории, так и эксперимента. Работа не с самим объектом (явлением, процессом), а с его моделью дает возможность безболезненно, относительно быстро и без существенных затрат исследовать его свойства и поведение в любых мыслимых ситуациях (преимущества теории). В то же время вычислительные (компьютерные, симуляционные, имитационные) эксперименты с моделями объектов позволяют, опираясь на мощь современных вычислительных методов и технических инструментов информатики, подробно и глубоко изучать объекты в достаточной полноте, недоступной чисто теоретическим подходам (преимущества эксперимента). Неудивительно, что методология математического моделирования бурно развивается, охватывая все новые сферы — от разработки технических систем и управления ими до анализа сложнейших экономических и социальных процессов.

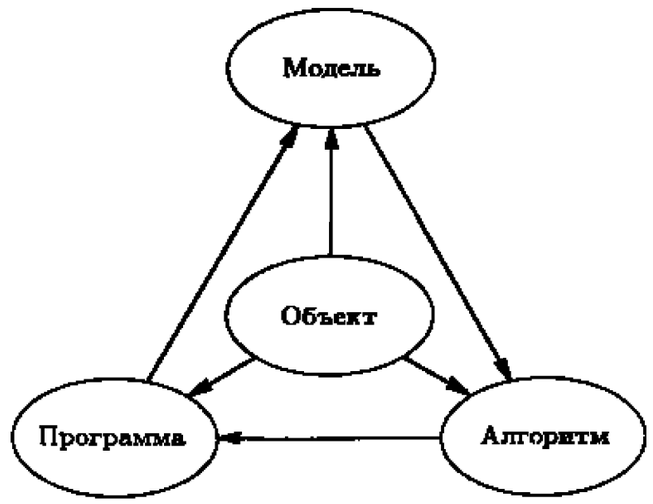
Элементы математического моделирования использовались с самого начала появления точных наук, и не случайно, что некоторые методы вычислений носят имена таких корифеев науки, как Ньютон и Эйлер, а слово «алгоритм» происходит от имени средневекового арабского ученого Аль-Хорезми. Второе «рождение» этой методологии пришлось на конец 40-х—начало 50-х годов XX века и было обусловлено по крайней мере двумя причинами. Первая из них — появление ЭВМ (компьютеров), хотя и скромных по нынешним меркам, но тем не менее избавивших ученых от огромной по объему рутинной вычислительной работы. Вторая — беспрецедентный социальный заказ — выполнение национальных программ СССР и США по созданию ракетно-ядерного щита, которые не могли быть реализованы традиционными методами. Математическое моделирование справилось с этой задачей: ядерные взрывы и полеты ракет и спутников были предварительно «осуществлены» в недрах ЭВМ с помощью математических моделей и лишь затем претворены на практике. Этот успех во многом определил дальнейшие достижения методологии, без применения которой в развитых странах ни один крупномасштабный технологический, экологический или экономический проект теперь всерьез не

рассматривается (сказанное справедливо и по отношению к некоторым социально-политическим проектам).

Сейчас математическое моделирование вступает в третий принципиально важный этап своего развития, «встраиваясь» в структуры так называемого *информационного общества*. Впечатляющий прогресс средств переработки, передачи и хранения информации отвечает мировым тенденциям к усложнению и взаимному проникновению различных сфер человеческой деятельности. Без владения информационными «ресурсами» нельзя и думать о решении все более укрупняющихся и все более разнообразных проблем, стоящих перед мировым сообществом. Однако информация как таковая зачастую мало что дает для анализа и прогноза, для принятия решений и контроля за их исполнением. Нужны надежные способы переработки информационного «сырья» в готовый «продукт», т. е. в точное знание. История методологии математического моделирования убеждает: она может и должна быть *интеллектуальным ядром* информационных технологий, всего процесса информатизации общества.

Технические, экологические, экономические и иные системы, изучаемые современной наукой, больше не поддаются исследованию (в нужной полноте и точности) обычными теоретическими методами. Прямой натурный эксперимент над ними долог, дорог, часто либо опасен, либо попросту невозможен, так как многие из этих систем существуют в «единственном экземпляре». Цена ошибок и просчетов в обращении с ними недопустимо высока. Поэтому *математическое* (*шире — информационное*) *моделирование является неизбежной составляющей научно-технического прогресса.*

Сама постановка вопроса о математическом моделировании какого-либо объекта порождает четкий план действий. Его можно условно разбить на три этапа: *модель — алгоритм — программа* (см. схему).



На первом этапе выбирается (или строится) «эквивалент» объекта, отражающий в математической форме важнейшие его свойства — законы, которым он подчиняется, связи, присущие составляющим его

частям, и т. д. Математическая модель (или ее фрагменты) исследуется теоретическими методами, что позволяет получить важные предварительные знания об объекте.

Второй этап — выбор (или разработка) алгоритма для реализации модели на компьютере. Модель представляется в форме, удобной для применения численных методов, определяется последовательность вычислительных и логических операций, которые нужно произвести, чтобы найти искомые величины с заданной точностью. Вычислительные алгоритмы должны не искажать основные свойства модели и, следовательно, исходного объекта, быть экономичными и адаптирующимися к особенностям решаемых задач и используемых компьютеров.

На третьем этапе создаются программы, «переводящие» модель и алгоритм на доступный компьютеру язык. К ним также предъявляются требования экономичности и адаптивности. Их можно назвать «электронным» эквивалентом изучаемого объекта, уже пригодным для непосредственного испытания на «экспериментальной установке» — компьютере.

Создав *триаду* «модель—алгоритм—программа», исследователь получает в руки универсальный, гибкий и недорогой инструмент, который вначале отлаживается, тестируется в «пробных» вычислительных экспериментах. После того как *адекватность* (достаточное соответствие) триады исходному объекту удостоверена, с моделью проводятся разнообразные и подробные «опыты», дающие все требуемые качественные и количественные свойства и характеристики объекта. Процесс моделирования сопровождается улучшением и уточнением, по мере необходимости, всех звеньев триады.

Будучи методологией, математическое моделирование не подменяет собой математику, физику, биологию и другие научные дисциплины, не конкурирует с ними. Наоборот, трудно переоценить его синтезирующую роль. Создание и применение триады невозможно без опоры на самые разные методы и подходы — от качественного анализа нелинейных моделей до современных языков программирования. Оно дает новые дополнительные стимулы самым разным направлениям науки.

Рассматривая вопрос шире, напомним, что моделирование присутствует почти во всех видах творческой активности людей различных «специальностей» — исследователей и предпринимателей, политиков и военачальников. Привнесение в эти сферы точного знания помогает ограничить интуитивное умозрительное «моделирование», расширяет поле приложений рациональных методов. Конечно же, математическое моделирование плодотворно лишь при выполнении хорошо известных профессиональных требований: четкая формулировка основных понятий и предположений, апостериорный анализ адекватности используемых моделей, гарантированная точность вычислительных алгоритмов и т. д. Если же говорить о моделировании систем с участием «человеческого фактора», т. е. *трудноформализуемых объектов*, то к этим требованиям необходимо добавить аккуратное разграничение математических и житейских терминов (звучащих одинаково, но имеющих разный смысл), осторожное применение уже готового математического аппарата к изучению явлений и процессов (предпочтителен путь «от задачи к методу», а не наоборот) и ряд других.

Решая проблемы информационного общества, было бы наивно уповать только на мощь компьютеров и иных средств информатики. Постоянное совершенствование триады математического моделирования и ее внедрение в современные информационно-моделирующие системы — *методологический императив*. Лишь его выполнение дает возможность получать так нужную нам высокотехнологичную, конкурентоспособную и разнообразную материальную и интеллектуальную продукцию.

Различным аспектам математического моделирования посвящено немало, хотя явно недостаточно, хороших и разных книг. При написании данной книги авторы ставили перед собой цель отобрать и изложить подходы к построению и анализу математических моделей, общие для различных областей знания, не зависящие от конкретной специфики. Окружающий людей мир един, и исследователи эффективно используют этот дар природы, выражающийся в том числе в *универсальности* математических моделей. Конечно, содержание книги связано в определенной мере с личным опытом авторов. Тем не менее, придерживаясь сформулированной выше точки зрения, нам было легче расширять рамки изложения и демонстрировать широкие возможности математического моделирования — от механики до социальных наук. Этим, как полагают авторы, данная книга и отличается от книг наших коллег.

Что касается стиля книги, то мы стремились избегать громоздких и строгих процедур (которые заинтересовавшийся читатель может найти в специальных изданиях), а обращали основное внимание на описание идей и соответствующих им примеров. Поэтому книга содержит большое количество иллюстраций и упражнений, ее материал без особых затруднений может быть использован для формирования учебных курсов по различным конкретным направлениям математического моделирования.

Небольшой объем книги не позволил осветить ряд важных тем; хотелось бы более подробно рассмотреть подходы к построению дискретных моделей и численные методы, вопросы «оснащения» моделей и их идентификации, их синтеза и декомпозиции и т. д. По этой же причине список литературы включает лишь необходимый минимум — это либо работы, результаты которых непосредственно отражены в тексте, либо ключевые книги, где можно найти дальнейшие ссылки.

Подбор материала и манера его изложения отвечают концепциям отечественной школы математического моделирования, всегда находившейся на мировом уровне и широко трактующей данную методологию. Поэтому мы не можем назвать всех своих коллег, общение с которыми так или иначе повлияло на содержание книги. Авторы считают своим приятным долгом поблагодарить доктора физико-математических наук Н. В. Змитренко, взявшего на себя нелегкий труд по прочтению рукописи и сделавшего ряд ценных замечаний, докторов физико-математических наук Б. Н. Четверушкина, В. Ф. Тишкина и академика РАН А. А. Петрова за предоставлен-

ные материалы и плодотворные дискуссии по актуальным проблемам математического моделирования. Авторы признательны также доктору физико-математических наук В. Я. Карпову, кандидату физико-математических наук А. Е. Королеву, кандидату технических наук В. А. Овсянникову, инженеру В. И. Зелепневу, экономисту О. А. Бусыгину за полезное обсуждение некоторых специальных вопросов. Мы благодарны также сотрудникам Института математического моделирования РАН Н. Г. Сиротенко и А. С. Болдареву, профессионально изготовившим оригинал-макет книги.

Надеемся, что наша книга будет интересна и полезна как начинающим, так и опытным исследователям и педагогам, использующим методы математического моделирования и вычислительного эксперимента в своей научной деятельности.

Москва, июнь 1997 г. *А. А. Самарский, А. П. Михайлов*

ГЛАВА I

ПРОСТЕЙШИЕ МАТЕМАТИЧЕСКИЕ МОДЕЛИ И ОСНОВНЫЕ ПОНЯТИЯ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ

§ 1. Элементарные математические модели

Рассмотрим некоторые подходы к построению простейших математических моделей, иллюстрирующие применение фундаментальных законов природы, вариационных принципов, аналогий, иерархических цепочек. Несмотря на простоту, привлекаемый материал даст возможность начать обсуждение таких понятий, как адекватность моделей, их «оснащение», нелинейность, численная реализация и ряда других принципиальных вопросов математического моделирования.

1. Фундаментальные законы природы. Наиболее распространенный метод построения моделей состоит в применении фундаментальных законов природы к конкретной ситуации. Эти законы общепризнаны, многократно подтверждены опытом, служат основой множества научно-технических достижений. Поэтому их обоснованность не вызывает сомнений, что, помимо всего прочего, обеспечивает исследователю мощную психологическую поддержку. На первый план выдвигаются вопросы, связанные с тем, какой закон (законы) следует применять в данном случае и как это делать.

а) *Сохранение энергии.* Этот закон известен почти двести лет и занимает, пожалуй, наиболее почетное место среди великих законов природы. Полагаясь на него, эксперт по баллистике, желающий быстро определить скорость револьверной пули и не имеющий поблизости специальной лаборатории, может воспользоваться относительно простым устройством типа маятника — груза, подвешенного на легком жестком и свободно вращающемся стержне (рис. 1). Пуля, застрявшая в грузе, сообщит системе «пуля—груз» свою кинетическую энергию, которая в момент наибольшего отклонения стержня от вертикали полностью перейдет в потенциальную энергию системы. Эти трансформации описываются цепочкой равенств

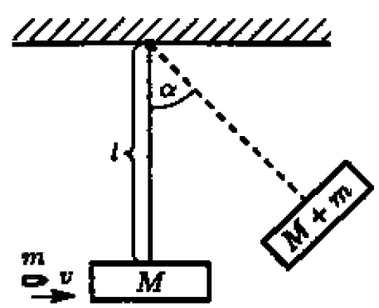


Рис. 1

$$\frac{mv^2}{2} = (M+m)\frac{V^2}{2} = (M+m)\,gl\,(1-\cos\alpha).$$

Здесь $mv^2/2$ — кинетическая энергия пули массы m, имеющей скорость v, M — масса груза, V — скорость системы «пуля—груз» сразу после столкновения, g — ускорение свободного падения, l — длина стержня, α — угол наибольшего отклонения. Искомая скорость определяется формулой

$$v = \sqrt{\frac{2(M+m)\,gl\,(1-\cos\alpha)}{m}}, \qquad (1)$$

которая будет вполне точной, если не учитываемые нами потери энергии на разогрев пули и груза, на преодоление сопротивления воздуха, разгон стержня и т. д. невелики. Это, на первый взгляд, разумное рассуждение на самом деле неверно. Процессы, происходящие при «слипании» пули и маятника, уже не являются чисто механическими. Поэтому примененный для вычисления величины V закон сохранения механической энергии несправедлив: сохраняется полная, а не механическая энергия системы. Он дает лишь нижнюю границу для оценки скорости пули (для правильного решения этой простой задачи надо воспользоваться также законом сохранения импульса — см. упр. 1).

Сходные рассуждения может применить и инженер для оценки времени t_k сверления слоя металла толщины L лазером с мощностью W, излучение которого перпендикулярно поверхности материала (рис. 2). Если энергия лазера полностью идет на испарение столбика

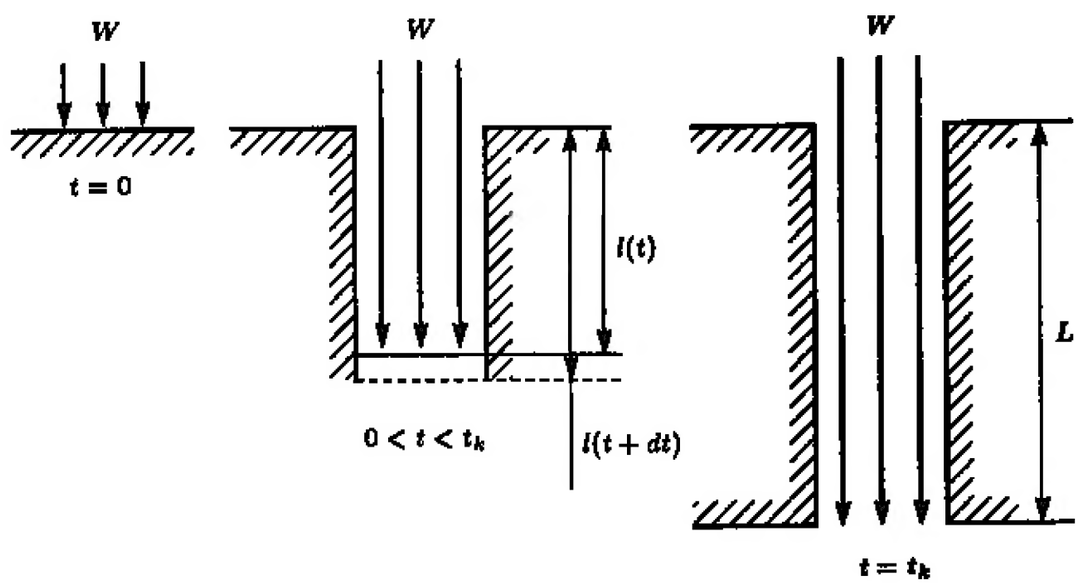


Рис. 2. Начальная, промежуточная и конечная стадии сверления металла лазером

металла массы $LS\rho$ (S — облучаемая площадь, LS — объем столбика, ρ — плотность вещества), то закон сохранения энергии выражается равенством

$$E_0 = Wt_k = hLS\rho, \qquad (2)$$

где h — энергия, требуемая для испарения единицы массы. Величина h имеет составную структуру: $h = (T_{пл} - T)\,h_1 + h_2 + h_3$, поскольку

материал необходимо последовательно нагреть до температуры плавления $T_{пл}$, а затем расплавить и превратить в пар (T — исходная температура, h_1 — удельная теплоемкость, h_2 и h_3 — соответственно удельная теплота плавления и парообразования).

Изменение глубины выемки $l(t)$ со временем определяется из детального баланса энергии в промежутке времени от t до $t + dt$. На испаренную за это время массу

$$[l(t+dt) - l(t)]\,S\rho = dl\,S\rho$$

тратится энергия $dl\,hS\rho$, равная энергии $W\,dt$, сообщаемой веществу лазером:

$$dl\,hS\rho = W\,dt,$$

откуда получается дифференциальное уравнение

$$\frac{dl}{dt} = \frac{W}{hS\rho}.$$

Его интегрирование (с учетом того, что начальная глубина выемки равна нулю) дает

$$l(t) = \frac{W}{hS\rho}\,t = \frac{E(t)}{hS\rho}, \qquad (3)$$

где $E(t)$ — вся энергия, выделенная лазером к моменту времени t. Следовательно, глубина выемки пропорциональна затраченной энергии (причем величина t_k, когда $l(t_k) = L$, совпадает с вычисленной по формуле (2)).

В действительности процесс сверления гораздо сложнее рассмотренной схемы — энергия тратится на нагрев вещества, на удаление паров из выемки, которая может иметь неправильную форму, и т. д. Поэтому уверенность в правильности предложенного математического описания значительно меньше, чем в случае с пулей. Вопрос о соответствии объекта и его модели — один из центральных в математическом моделировании, и в дальнейшем мы будем неоднократно к нему возвращаться.

б) *Сохранение материи.* Именно этим соображением руководствуется школьник, решающий задачу о заполнении бассейна водой, втекающей и вытекающей из двух труб. Конечно же, область применения этого закона несравненно шире.

Пусть, например, имеется небольшое количество радиоактивного вещества (урана), окруженного толстым слоем «обычного» материала (свинца), — ситуация типичная либо при хранении делящихся материалов, либо при их использовании в энергетике (рис. 3). Под словом «небольшой» подразумевается упрощающее обстоятельство, а именно то, что все продукты распада, не испытывая столкновений с атомами вещества, беспрепятственно покидают область I. Другими словами, длина свободного пробега продуктов распада λ_I в первом веществе значительно больше характерных размеров самого материа-

ла $L_{\rm I}$, т. е. $\lambda_{\rm I} \gg L_{\rm I}$. Слова «толстый слой» означают, что в согласии с целями хранения продукты деления полностью поглощаются в области II. Это гарантируется при выполнении противоположного условия $\lambda_{\rm II} \ll L_{\rm II}$, где $\lambda_{\rm II}$ — длина пробега продуктов распада во втором веществе, $L_{\rm II}$ — его характерный размер.

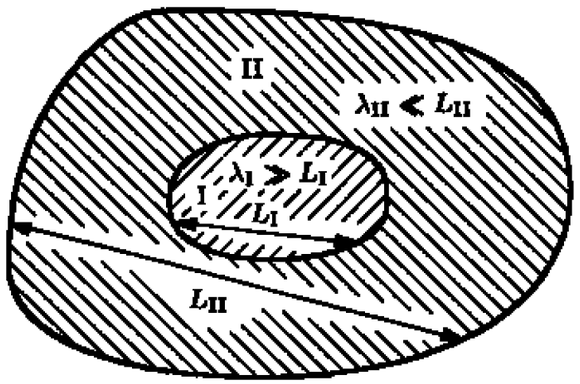


Рис. 3

Итак, все, что вылетает из области I, поглощается в области II, и суммарная масса обоих веществ со временем не меняется. Это и есть закон сохранения материи, примененный к данной ситуации. Если в начальный момент времени $t = 0$ массы веществ были равны $M_{\rm I}(0)$ и $M_{\rm II}(0)$, то в любой момент времени справедлив баланс

$$M_{\rm I}(0) + M_{\rm II}(0) = M_{\rm I}(t) + M_{\rm II}(t). \tag{4}$$

Одного уравнения (4), очевидно, недостаточно для определения текущих значений двух масс — $M_{\rm I}(t)$ и $M_{\rm II}(t)$. Для замыкания математической формулировки необходимо привлечь дополнительное соображение о характере распада. Оно гласит, что скорость распада (число атомов, распадающихся в единицу времени) пропорционально общему числу атомов радиоактивного вещества. За небольшое время dt между моментами t и $t + dt$ всего распадется

$$N_{\rm I}(t + dt) - N_{\rm I}(t) = -\alpha N_{\rm I}(t + \xi\, dt), \qquad \alpha > 0, \quad 0 < \xi < 1,$$

атомов. Здесь вторично использован закон сохранения вещества, но применительно не ко всему процессу, а к отрезку времени dt. В этом уравнении, описывающем баланс атомов, в правой части стоит знак минус (вещество убывает), а величина $N_{\rm I}(t + \xi\, dt)$ отвечает некоторому среднему значению числа атомов за рассматриваемое время. Перепишем его в дифференциальной форме:

$$\frac{dN_{\rm I}(t)}{dt} = -\alpha N_{\rm I}(t).$$

Учитывая, что $M_{\rm I}(t) = \mu_{\rm I} N_{\rm I}(t)$, где $\mu_{\rm I}$ — атомный вес вещества I, получаем

$$\frac{dM_{\rm I}(t)}{dt} = -\alpha M_{\rm I}(t). \tag{5}$$

При самопроизвольной радиоактивности любой атом имеет некоторую не зависящую от состояния окружающего вещества вероятность распада. Поэтому чем больше (меньше) самого радиоактивного вещества, тем больше (меньше) выделяется продуктов распада в единицу времени. Коэффициент пропорциональности $\alpha > 0$ (*постоянная распада*) определяется конкретным веществом.

Уравнения (4), (5) вместе с условиями $\lambda_{\rm I} \gg L_{\rm I}$, $\lambda_{\rm II} \ll L_{\rm II}$, а также величинами α, $M_{\rm I}(0)$, $M_{\rm II}(0)$ и составляют математическую модель рассматриваемого объекта.

Интегрируя (5), получаем, что масса делящегося материала убывает по экспоненциальному закону

$$M_{\rm I}(t) = M_{\rm I}(0)\, e^{-\alpha t},$$

и при $t \to \infty$ в области I вещество полностью исчезает.

Так как суммарная масса в соответствии с (4) остается постоянной, то в области II количество вещества растет:

$$M_{\rm II}(t) = M_{\rm II}(0) + M_{\rm I}(0) - M_{\rm I}(0)\, e^{-\alpha t} = M_{\rm II}(0) + M_{\rm I}(0)\left(1 - e^{-\alpha t}\right),$$

и при $t \to \infty$ продукты распада полностью переходят из области I в область II

в) *Сохранение импульса.* Неподвижно стоящая в безветренную погоду на поверхности озера лодка начнет двигаться вперед, если сделать несколько шагов от ее носа к корме. Так проявляет себя закон сохранения импульса, утверждающий: полный импульс системы, не испытывающей действия внешних сил, сохраняется. На передвижение гребца лодка реагирует смещением в противоположную сторону.

Принцип реактивного движения положен в основу многих замечательных технических устройств, например, ракеты, выводящей на орбиту вокруг Земли искусственный спутник, для чего ей требуется развить скорость примерно 8 км/с. Простейшая математическая модель движения ракеты получается из закона сохранения импульса в пренебрежении сопротивлением воздуха, гравитацией и другими силами, исключая, конечно, тягу реактивных двигателей.

Пусть продукты сгорания ракетного топлива покидают расположенные в кормовой части выхлопные сопла со скоростью u (для современных топлив величина u равна 3–5 км/с). За малый промежуток времени dt между моментами t и $t + dt$ часть топлива выгорела, и масса ракеты изменилась на величину dm. Изменился также импульс ракеты, однако суммарный импульс системы «ракета плюс продукты сгорания» остался тем же, что и в момент t, т. е.

$$m(t)\, v(t) = m(t + dt)\, v(t + dt) - dm\, [v(t + \xi\, dt) - u],$$

где $v(t)$ — скорость ракеты, $v(t + \xi\, dt) - u$ $(0 < \xi < 1)$ — средняя за промежуток dt скорость истекающих из сопел газов (обе скорости берутся относительно Земли). Первый член в правой части этого равенства — импульс ракеты в момент $t + dt$, второй — импульс, переданный истекающим газом за время dt.

Учитывая, что $m(t + dt) = m(t) + (dm/dt)\, dt + O(dt^2)$, закон сохранения импульса можно переписать в виде дифференциального уравнения

$$m\,\frac{dv}{dt} = -\,\frac{dm}{dt}\, u,$$

в котором член $-(dm/dt)\,u$, очевидно, не что иное, как сила тяги ракетных двигателей, и которое, будучи преобразованным к виду

$$\frac{dv}{dt} = -u\,\frac{d(\ln m)}{dt},$$

легко интегрируется:

$$v(t) = v_0 + u\ln\left(\frac{m_0}{m(t)}\right),$$

где v_0, m_0 — соответственно скорость и масса ракеты в момент $t = 0$. Если $v_0 = 0$, то максимальная скорость ракеты, достигаемая при полном сгорании топлива, равна

$$v = u\ln\left(\frac{m_0}{m_p + m_s}\right). \tag{6}$$

Здесь m_p — полезная масса (масса спутника), m_s — структурная масса (масса собственно ракетной конструкции — топливных баков, двигателей, систем управления и т. д.).

Простая формула Циолковского (6) позволяет сделать фундаментальный вывод о конструкции ракеты для космических полетов. Введем величину $\lambda = \dfrac{m_s}{m_0 - m_p}$, которая характеризует при $m_p = 0$ отношение структурной и начальной масс ракеты. Тогда для практически реальных значений $\lambda = 0{,}1$, $u = 3$ км/с получаем при $m_p = 0$

$$v = u\ln(1/\lambda) = 7 \text{ км/с}.$$

Отсюда следует, что даже в самой идеальной ситуации (полезная масса равна нулю, отсутствуют гравитация и сопротивление воздуха и т. д.) ракета рассматриваемого типа не способна достичь первой космической скорости. Тем самым необходимо использовать многоступенчатые ракеты — вывод, к которому пришли основоположники космонавтики.

Данный пример иллюстрирует также своего рода принцип «наибольшего благоприятствия», часто используемый на начальной стадии математического моделирования сложных объектов: если объект, поставленный в наилучшие условия, не в состоянии достичь требуемых характеристик, то надо изменить сам подход к объекту либо смягчить требования к нему; если же требования в принципе достижимы, то следующие шаги связаны с исследованием влияния на объект дополнительных осложняющих факторов.

2. Вариационные принципы. Еще один подход к построению моделей, по своей широте и универсальности сопоставимый с возможностями, даваемыми фундаментальными законами, состоит в применении так называемых *вариационных принципов*. Они представляют собой весьма общие утверждения о рассматриваемом объекте (системе, явлении) и гласят, что из всех возможных вариантов его поведения (движения, эволюции) выбираются лишь те, которые удовлетворяют определенному условию. Обычно согласно этому условию некоторая

связанная с объектом величина достигает экстремального значения при его переходе из одного состояния в другое.

Допустим, автомобиль, движущийся с постоянной скоростью v, должен попасть из точки A в точку B и при этом коснуться некоторой прямой линии C (рис. 4). Водитель автомобиля очень торопится и выбирает из множества траекторий путь, требующий минимальных

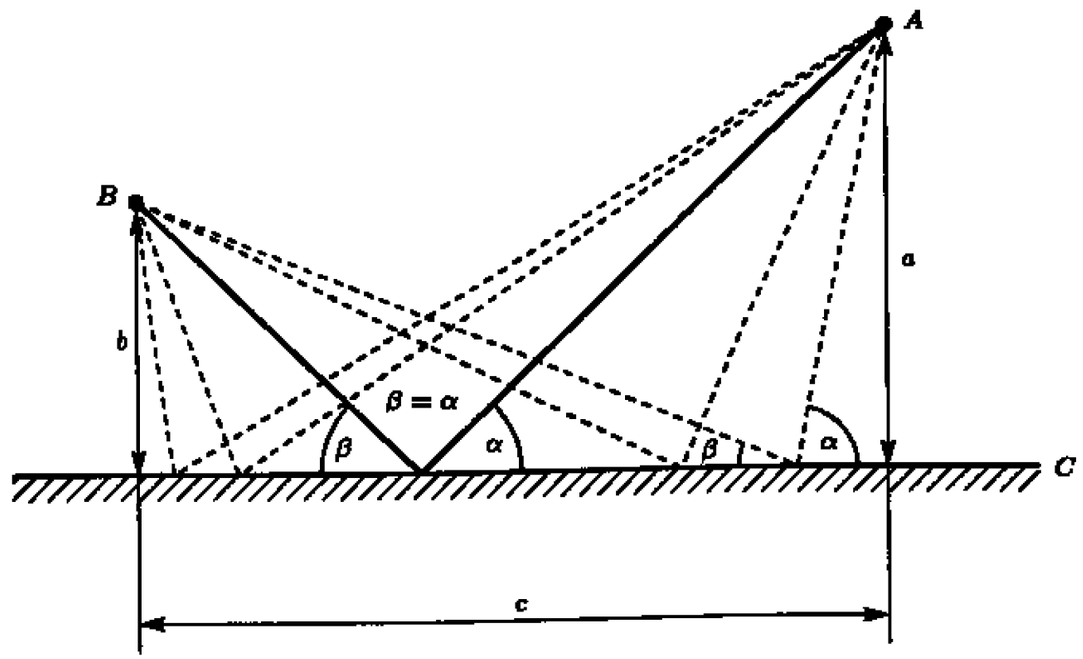


Рис. 4. Различные траектории движения из точки A в точку B с касанием прямой C. Жирной линией выделен быстрейший путь

затрат времени. Представим затраченное время как функцию величины α — угла между прямой и отрезком пути от точки A до прямой:

$$t(\alpha) = \frac{a}{v\sin\alpha} + \frac{b}{v\sin\beta(\alpha)}.$$

Здесь a и b — длины перпендикуляров, опущенных из точек A и B на прямую, $\beta(\alpha)$ — угол между прямой и отрезком пути из точки касания до точки B.

Условие экстремальности $t(\alpha)$ по аргументу α означает, что

$$\left.\frac{dt(\alpha)}{d\alpha}\right|_{\alpha=\alpha_{\text{ext}}} = 0,$$

или

$$\frac{a\cos\alpha}{\sin^2\alpha} + \frac{b\cos\beta(\alpha)}{\sin^2\beta(\alpha)}\,\frac{d\beta}{d\alpha} = 0. \tag{7}$$

Для любых значений α справедливо равенство

$$c = \frac{a}{\operatorname{tg}\alpha} + \frac{b}{\operatorname{tg}\beta(\alpha)},$$

где c — расстояние между проекциями точек A и B на прямую (одинаковое для всех траекторий). Дифференцируя его, получаем соотношение

$$\frac{a}{\sin^2\alpha} + \frac{b}{\sin^2\beta(\alpha)}\frac{d\beta}{d\alpha} = 0, \tag{8}$$

которое вместе с условием минимальности (7) означает

$$\cos\alpha = \cos\beta(\alpha),$$

т. е. равенство углов α и β (см. упр. 5).

Далее нетрудно найти сами значения $\alpha_{\min}, t_{\min}$ через заданные величины a, b, c. Однако сейчас для нас важно другое — условие минимальных затрат времени привело к выбору соответствующей траектории по правилу «угол падения равен углу отражения». Но ведь такому закону подчиняется и ход светового луча, попадающего на отражающую поверхность! Может быть, и в общем случае лучи света движутся по траекториям, обеспечивающим быстрейшее попадание сигнала из одной точки в другую? Да, именно так и происходит согласно известному вариационному принципу Ферма, опираясь на который, можно получить все основные законы геометрической оптики.

Покажем это, рассмотрев преломление лучей на границе двух сред (рис. 5). Свет, выходящий из точки A, движется в первой среде со скоростью v_a, преломляется и, переходя через линию раздела, двигается

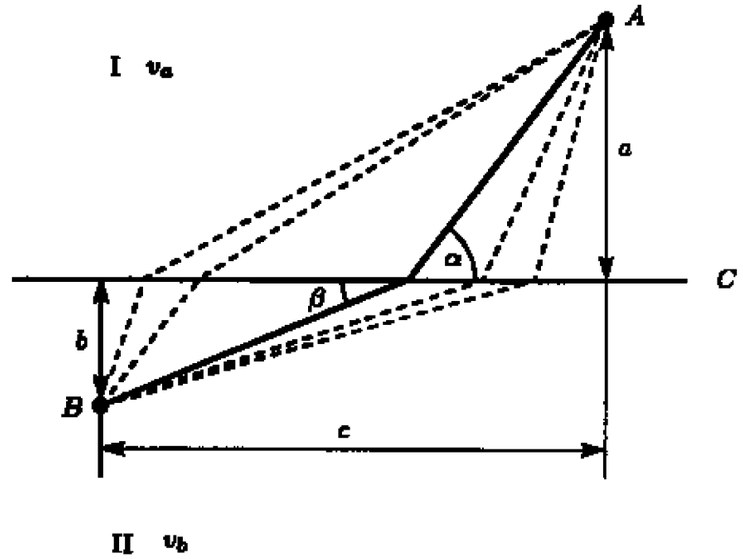


Рис. 5. Возможные траектории световых лучей, идущих из точки A в точку B и преломляющихся на линии C — границе раздела двух сред. Жирной линией выделена траектория, отвечающая закону преломления $\cos\alpha/\cos\beta = v_a/v_b$

во второй среде со скоростью v_b и попадает в точку B. Если α — угол падения луча, а $\beta(\alpha)$ — угол его преломления, то время прохождения

из A в B равно

$$t(\alpha) = \frac{a}{v_a\sin\alpha} + \frac{b}{v_b\sin\beta(\alpha)}.$$

Условие минимальности $t(\alpha)$ записывается в виде (ср. с (7))

$$\frac{a\cos\alpha}{v_a\sin\alpha} + \frac{b\cos\alpha}{v_b\sin\beta(\alpha)}\frac{d\beta}{d\alpha} = 0,$$

а продифференцированное по α условие постоянства величины c по-прежнему выражается формулой (8). Здесь величины a, b, c имеют тот же смысл, что и в предыдущем случае. Исключая из последней формулы производную $d\beta/d\alpha$, приходим к равенству

$$\frac{\cos\alpha}{\cos\beta} = \frac{v_a}{v_b}, \tag{9}$$

т. е. к известному закону преломления света.

Сформулированные применительно к какому-либо классу явлений вариационные принципы позволяют единообразно строить соответствующие математические модели. Их универсальность выражается также в том, что, используя их, можно в определенной степени отвлекаться от конкретной природы процесса. Так, водитель автомобиля, следующий принципу «минимального времени» и желающий попасть из точки A, находящейся на песчаной почве (одна скорость), в точку B, расположенную на травянистом лугу (другая скорость), обязан поехать не по прямой, соединяющей A и B, а по ломанной траектории, сделав необходимое «преломление» на линии, разделяющей песок и траву.

3. Применение аналогий при построении моделей. В огромном числе случаев при попытке построить модель какого-либо объекта либо невозможно прямо указать фундаментальные законы или вариационные принципы, которым он подчиняется, либо, с точки зрения наших сегодняшних знаний, вообще нет уверенности в существовании подобных законов, допускающих математическую формулировку. Одним из плодотворных подходов к такого рода объектам является использование аналогий с уже изученными явлениями. Что, казалось бы, общего между радиоактивным распадом и динамикой популяций, в частности изменением численности населения нашей планеты? Однако на простейшем уровне такая аналогия вполне просматривается, о чем свидетельствует одна из простейших моделей популяций, называемая *моделью Мальтуса*. В ее основу положено простое утверждение - скорость изменения населения со временем t пропорциональна его текущей численности $N(t)$, умноженной на сумму коэффициентов рождаемости $\alpha(t) \geqslant 0$ и смертности $\beta(t) \leqslant 0$. В результате приходим к уравнению

$$\frac{dN(t)}{dt} = [\alpha(t) - \beta(t)]\,N(t), \tag{10}$$

весьма похожему на уравнение радиоактивного распада и совпадающего с ним при $\alpha < \beta$ (если α и β постоянные). Это неудивительно,

так как при их выводе использовались одинаковые соображения. Интегрирование уравнения (10) даст

$$N(t) = N(0)\exp\left(\int_{t_0}^{t}[\alpha(t)-\beta(t)]\,dt\right),$$

где $N(0) = N(t = t_0)$ — начальная численность.

На рис. 6 приведены графики функции $N(t)$ при постоянных α и β (разным подобным друг другу кривым соответствуют разные t_0 — значения времени начала процесса). При $\alpha = \beta$ численность остается постоянной, т. е. в этом случае решением уравнения является равновесная величина $N(t) = N(0)$. Равновесие между рождаемостью и

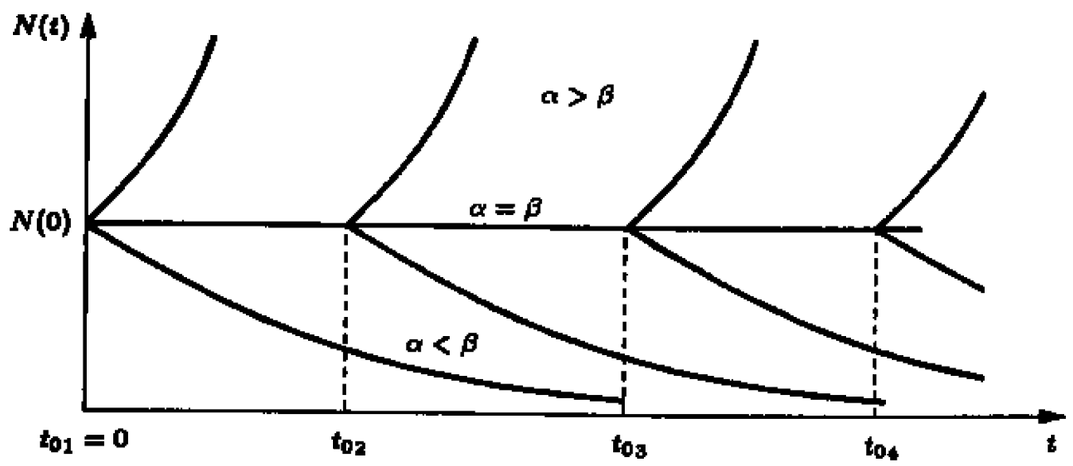


Рис. 6. Изменение численности популяции со временем в модели Мальтуса

смертностью неустойчиво в том смысле, что даже небольшое нарушение равенства $\alpha = \beta$ приводит с течением времени ко все большему отклонению функции $N(t)$ от равновесного значения $N(0)$. При $\alpha < \beta$ численность населения убывает и стремится к нулю при $t \to \infty$, а при $\alpha > \beta$ растет по некоторому экспоненциальному закону, обращаясь в бесконечность при $t \to \infty$. Последнее обстоятельство и послужило основанием для опасений Мальтуса о грядущем перенаселении Земли со всеми вытекающими отсюда последствиями.

Как в данном примере, так и в ряде рассмотренных выше случаев можно указать немало очевидных ограничений применимости построенной модели. Конечно же, сложнейший процесс изменения численности населения, зависящий к тому же от сознательного вмешательства самих людей, не может описываться какими-либо простыми закономерностями. Даже в идеальном случае изолированной биологической популяции предложенная модель не отвечает реальности в полной мере хотя бы из-за ограниченности ресурсов, необходимых для ее существования.

Сделанное замечание тем не менее нисколько не умаляет роли аналогий в построении математических моделей очень сложных явлений.

Применение аналогий основано на одном из важнейших свойств моделей — их универсальности, т. е. их приложимости к объектам принципиально различной природы. Так, предположения типа «скорость изменения величины пропорциональна значению самой величины (или некоторой функции от нее)» широко используются в далеких друг от друга областях знаний.

4. Иерархический подход к получению моделей. Лишь в редких случаях бывает удобным и оправданным построение математических моделей даже относительно простых объектов сразу во всей полноте, с учетом всех факторов, существенных для его поведения. Поэтому естествен подход, реализующий принцип «от простого — к сложному», когда следующий шаг делается после достаточно подробного изучения не очень сложной модели. При этом возникает цепочка (*иерархия*) все более полных моделей, каждая из которых обобщает предыдущие, включая их в качестве частного случая.

Построим такую иерархическую цепочку на примере модели многоступенчатой ракеты. Как было установлено в конце п. 1, реальная одноступенчатая ракета неспособна развить первую космическую скорость. Причина этого — затраты горючего на разгон ненужной, отработавшей части структурной массы. Следовательно, при движении ракеты необходимо периодически избавляться от балласта. В практической конструкции это означает, что ракета состоит из нескольких ступеней, отбрасываемых по мере их использования.

Пусть m_i — общая масса i-й ступени, λm_i — соответствующая структурная масса (при этом масса топлива равна величине $(1-\lambda)m_i$), m_p — масса полезной нагрузки. Величины λ и скорость истечения газов u одинаковы для всех ступеней. Возьмем для определенности число ступеней $n = 3$. Начальная масса такой ракеты равна

$$m_0 = m_p + m_1 + m_2 + m_3.$$

Рассмотрим момент, когда израсходовано все топливо первой ступени и масса ракеты равна величине

$$m_p + \lambda m_1 + m_2 + m_3.$$

Тогда по формуле (6) первоначальной модели скорость ракеты равна

$$v_1 = u\ln\left(\frac{m_0}{m_p + \lambda m_1 + m_2 + m_3}\right).$$

После достижения скорости v_1 структурная масса λm_1 отбрасывается и включается вторая ступень. Масса ракеты в этот момент равна

$$m_p + m_2 + m_3.$$

Начиная с этого момента и до момента полного выгорания топлива второй ступени, ничто не мешает пользоваться уже построенной моделью, применив ее к рассматриваемому случаю. Все рассуждения о сохранении суммарного импульса и соответствующие выкладки остаются в силе (следует только учесть, что у ракеты уже есть начальная

скорость v_1). Тогда по формуле (6) после выгорания топлива во второй ступени ракета достигает скорости

$$v_2 = v_1 + u \ln\left(\frac{m_p + m_2 + m_3}{m_p + \lambda m_2 + m_3}\right).$$

Такие же рассуждения применимы и к третьей ступени ракеты. После отключения ее двигателей скорость ракеты равна

$$v_3 = v_2 + u \ln\left(\frac{m_p + m_3}{m_p + \lambda m_3}\right).$$

Эту цепочку нетрудно продолжить для любого числа ступеней и получить соответствующие формулы. В случае же $n = 3$ для окончательной скорости имеем

$$\frac{v_3}{u} = \ln\left\{\left(\frac{m_0}{m_p + \lambda m_1 + m_2 + m_3}\right)\left(\frac{m_p + m_2 + m_3}{m_p + \lambda m_2 + m_3}\right)\left(\frac{m_p + m_3}{m_p + \lambda m_3}\right)\right\},$$

или, вводя величины $\alpha_1 = \dfrac{m_0}{m_p + m_2 + m_3}$, $\alpha_2 = \dfrac{m_p + m_2 + m_3}{m_p + m_3}$, $\alpha_3 = \dfrac{m_p + m_3}{m_p}$, получаем

$$\frac{v_3}{u} = \ln\left\{\left(\frac{\alpha_1}{1 + \lambda(\alpha_1 - 1)}\right)\left(\frac{\alpha_2}{1 + \lambda(\alpha_2 - 1)}\right)\left(\frac{\alpha_3}{1 + \lambda(\alpha_3 - 1)}\right)\right\}.$$

Данное выражение симметрично по отношению к величинам $\alpha_1, \alpha_2, \alpha_3$, и нетрудно показать, что его максимум достигается в симметричном случае, т. е. при $\alpha_1 = \alpha_2 = \alpha_3 = \alpha$. При этом для $i = 3$

$$\alpha = \frac{1 - \lambda}{P - \lambda}, \qquad P = \exp\left(-\frac{v_3}{3u}\right).$$

Произведение $\alpha_1\alpha_2\alpha_3 = \alpha$ равно, как легко проверить, отношению m_0/m_p, или

$$\alpha^3 = \frac{m_0}{m_p} = \left(\frac{1 - \lambda}{P - \lambda}\right)^3.$$

Для многоступенчатой ракеты, аналогично, имеем

$$\frac{m_0}{m_p} = \left(\frac{1 - \lambda}{P - \lambda}\right)^n, \qquad P = \exp\left(-\frac{v_n}{nu}\right), \tag{11}$$

где n — число ступеней.

Проанализируем формулу (11). Примем $v_n = 10{,}5$ км/с, $\lambda = 0{,}1$. Тогда для $n = 2, 3, 4$ получаем $m_0 = 149\,m_p$, $m_0 = 77\,m_p$, $m_0 = 65\,m_p$ соответственно. Это значит, что двухступенчатая ракета пригодна для выведения на орбиту некоторой полезной массы (однако при одной тонне полезного груза необходимо иметь ракету весом 149 тонн). Переход к третьей ступени уменьшает массу ракеты почти в два раза (но, конечно же, усложняет ее конструкцию), а четырехступенчатая ракета не дает заметного выигрыша по сравнению с трехступенчатой.

Построение иерархической цепочки позволило относительно просто прийти к этим важным выводам. Иерархия математических моделей часто строится и по противоположному принципу «от сложного к простому». В этом случае реализуется путь «сверху вниз» — из достаточно общей и сложной модели при соответствующих упрощающих предположениях получается последовательность все более простых (но имеющих уменьшающуюся область применимости) моделей.

5. О нелинейности математических моделей. Простота рассмотренных выше моделей во многом связана с их *линейностью*. В математическом плане это важное понятие означает, что справедлив *принцип суперпозиции*, т. е. любая линейная комбинация решений (например, их сумма) также является решением задачи. Пользуясь принципом суперпозиции, нетрудно, найдя решение в каком-либо частном случае, построить решение в более общей ситуации. Поэтому о качественных свойствах общего случая можно судить по свойствам частного — различие между двумя решениями носит лишь количественный характер. Например, увеличение в два раза скорости истечения ракетного топлива ведет также к двукратному увеличению скорости ракеты, уменьшение угла падения светового луча на отражающую поверхность означает такое же изменение угла отражения и т. д. Другими словами, в случае линейных моделей отклик объекта на изменение каких-то условий пропорционален величине этого изменения.

Для *нелинейных явлений*, математические модели которых не подчиняются принципу суперпозиции, знание о поведении части объекта еще не гарантирует знания поведения всего объекта, а его отклик на изменение условий может качественно зависеть от величины этого изменения. Так, уменьшение угла падения луча света на границу раздела двух сред приводит к уменьшению угла преломления, но только до определенного предела. Если угол падения становится меньше критического (см. формулу (9)), то происходит качественное изменение — свет перестает проникать через границу раздела во вторую среду, если она менее плотная, чем первая. Тем самым преломление света — пример нелинейного процесса.

Большинство реальных процессов и соответствующих им математических моделей нелинейны. Линейные же модели отвечают весьма частным случаям и, как правило, служат лишь первым приближением к реальности. Например, популяционные модели сразу становятся нелинейными, если принять во внимание ограниченность доступных популяции ресурсов. При их выводе считается, что:

1) существует «равновесная» численность популяции N_p, которую может обеспечить окружающая среда;

2) скорость изменения численности популяции пропорциональна самой численности, умноженной (в отличие от модели Мальтуса) на величину ее отклонения от равновесного значения, т. е

$$\frac{dN}{dt} = \alpha\left(1 - \frac{N}{N_p}\right)N, \qquad \alpha > 0. \tag{12}$$

Член $(1 - N/N_p)$ в этом уравнении обеспечивает механизм «насыщения» численности — при $N < N_p$ $(N > N_p)$ скорость роста положительна (отрицательна) и стремится к нулю, если $N \to N_p$.

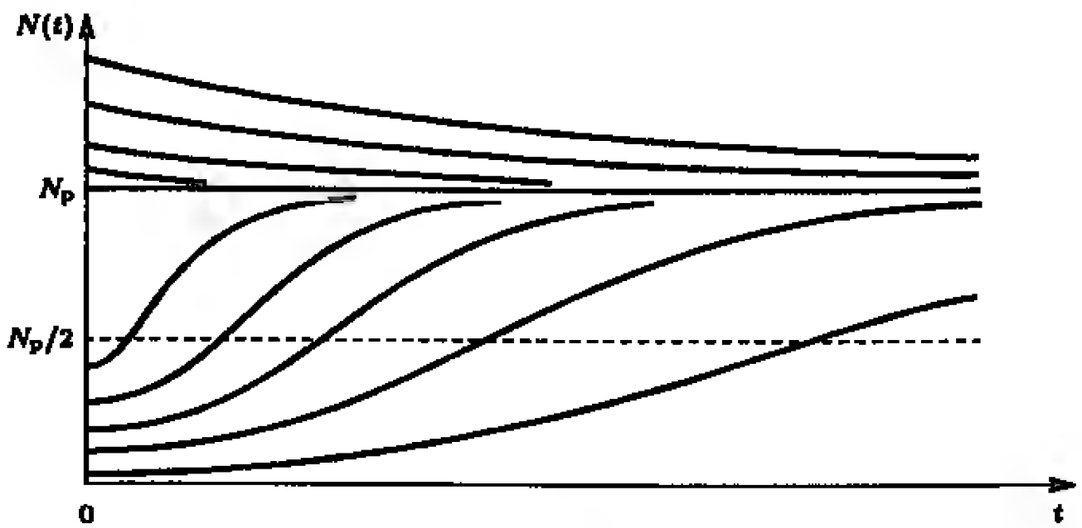


Рис. 7. Логистические кривые, соответствующие различным значениям начальной численности $N(0)$

Представляя уравнение (12) в виде

$$\frac{dN}{N_p - N} + \frac{dN}{N} = \alpha\, dt$$

и интегрируя его, получаем

$$-\ln(N_p - N) + \ln N = \alpha t + C.$$

Постоянная интегрирования определяется из условия $N(t = 0) = N(0)$, т. е. $C = \ln\left((N_p - N(0))^{-1}N(0)\right)$. В результате находим

$$N = N_p\,\frac{N(0)}{N_p - N(0)}\,e^{\alpha t} - N\,\frac{N(0)}{N_p - N(0)}\,e^{\alpha t},$$

или, в окончательном виде,

$$N(t) = \frac{N_p N(0)\,e^{\alpha t}}{N_p - N(0)\,(1 - e^{\alpha t})}.$$

Поведение функции $N(t)$ описывается так называемой *логистической кривой* (рис. 7). При любом $N(0)$ численность стремится к равновес-

ному значению N_p, причем тем медленней, чем величина $N(t)$ ближе к $N(0)$. Тем самым равновесие, в отличие от случая модели (10), устойчиво.

Логистическая модель более реалистично отражает динамику популяции в сравнении с моделью Мальтуса, но сама она с необходимостью становится нелинейной и поэтому более сложной. Заметим, что предположения о механизмах насыщения используются при построении многих моделей в различных областях знаний.

6. Предварительные выводы. Процесс построения моделей может быть условно разбит на следующие этапы.

1. Конструирование модели начинается со словесно-смыслового описания объекта или явления. Помимо сведений общего характера о природе объекта и целях его исследования эта стадия может содержать также некоторые предположения (невесомый стержень, толстый слой вещества, прямолинейное распространение световых лучей и т. д.). Данный этап можно назвать формулировкой предмодели.

2. Следующий этап — завершение идеализации объекта. Отбрасываются все факторы и эффекты, которые представляются не самыми существенными для его поведения. Например, при составлении баланса материи (п. 1б)) не учитывался, ввиду его малости, дефект масс, которым сопровождается радиоактивный распад. По возможности идеализирующие предположения записываются в математической форме (подобно условию $\lambda_{\rm I} \gg L_{\rm I}$ в п. 1б)), с тем чтобы их справедливость поддавалась количественному контролю.

3. После выполнения первых двух этапов можно переходить к выбору или формулировке закона (вариационного принципа, аналогии и т. п.), которому подчиняется объект, и его записи в математической форме. При небходимости используются дополнительные сведения об объекте, также записываемые математически (например, постоянство величины c для всех траекторий лучей света, вытекающее из геометрии задачи; п. 2). Следует иметь в виду, что даже для простых объектов выбор соответствующего закона отнюдь не тривиальная задача (см. упр. 1).

4. Завершает формулировку модели ее «оснащение». Например, необходимо задать сведения о начальном состоянии объекта (скорость ракеты и ее массу в момент $t = 0$) или иные его характеристики (величины l, g в п. 1а); α, $\lambda_{\rm I}$, $\lambda_{\rm II}$ в п. 1б); $\alpha(t)$ и $\beta(t)$ в п. 3), без знания которых невозможно определить поведение объекта. И, наконец, формулируется цель исследования модели (найти закон преломления света, достичь понимания закономерностей изменения популяции, определить требования к конструкции ракеты, запускающей спутник, и т. д.).

5. Построенная модель изучается всеми доступными исследователю методами, в том числе со взаимной проверкой различных подходов (см., например, упр. 4, 7). В отличие от рассматриваемых в § 1 простейших случаев, большинство моделей не поддаются чисто теоретическому анализу, и поэтому необходимо широко использовать вычислительные методы. Это обстоятельство особенно важно при изучении нелинейных объектов, так как их качественное поведение заранее, как правило, неизвестно.

6. В результате исследования модели не только достигается поставленная цель, но и должна быть установлена всеми возможными способами (сравнением с практикой, сопоставлением с другими подходами) ее адекватность — соответствие объекту и сформулированным предположениям. Неадекватная модель может дать результат, сколь угодно отличающийся от истинного (ср. формулу (1) и результат упр. 1), и должна быть либо отброшена, либо соответствующим образом модифицирована.

УПРАЖНЕНИЯ

1. В первой задаче п. 1, а) примените для нахождения величины V (скорости системы «пуля—груз» сразу после столкновения) не закон сохранения энергии, а закон сохранения импульса. Убедитесь, что для скорости пули v получается формула, дающая значение в $((M+m)/m)^{1/2}$ раз меньше, чем получающееся по формуле (1).

2. Пусть мощность лазера, сверлящего материал (п. 1, а)), зависит от времени, т. е. $W = W(t)$. Как изменится формула (3)? Остается ли в силе утверждение о том, что глубина выемки пропорциональна затраченной энергии?

3. Найдите момент времени, когда распадается последний атом радиоактивного вещества (п. 1, б)). Почему в модели (5) вещество распадается полностью лишь при $t \to \infty$?

4. Предположим, что в п. 1, в) рассматривается «идеальная» одноступенчатая ракета, у которой непрерывно отбрасывается отработавшая и ставшая ненужной часть структурной массы (к моменту полного сгорания топлива $m_s = 0$). Пользуясь законом сохранения импульса, покажите, что максимальная скорость такой ракеты определяется по формуле $v = (1-\lambda)\,u\ln(m_0/m_p)$. Сравните ее с формулой (6). Почему идеальная ракета может достичь любой скорости?

5. Проверьте с использованием (8), что условие (7) есть условие минимальности для величины $t(\alpha)$. Из рис. 5 определите, какая скорость больше — v_a или v_b? Пользуясь формулой (9), найдите, при каких углах α лучи света не проникают из среды a в среду b, т. е. когда реализуется эффект «полного внутреннего отражения света», используемый в ряде технических устройств.

6. Определите, как себя должна вести при больших t величина $r(t) = \alpha(t) - \beta(t) > 0$ в модели Мальтуса (10), чтобы численность популяции оставалась ограниченной при $t \to \infty$.

7. Перейдите в формуле (11) для многоступенчатой ракеты к пределу при $n \to \infty$ и убедитесь, что ее предельная скорость вычисляется по формуле для идеальной ракеты из упр. 4. Почему результаты совпадают?

8. Рассмотрите в логистической модели (12) малые отклонения от положения равновесия, т. е. ситуацию, когда решение имеет вид $N(t) = N_p + \delta N(t)$, где $|\delta N(t)| \ll N_p$. Покажите, что для величины $\delta N(t)$ в первом приближении справедлива линейная модель типа модели Мальтуса (10).

§ 2. Примеры моделей, получаемых из фундаментальных законов природы

По сравнению с п. 1 из § 1 более подробно и для более сложных объектов рассмотрим модели, следующие из законов Архимеда, Ньютона, Кулона и других хорошо известных законов. Обсудим некоторые свойства рассматриваемых объектов.

1. Траектория всплытия подводной лодки. Пусть подводная лодка, находящаяся в момент времени $t = 0$ на глубине H от поверхности моря и движущаяся с постоянной горизонтальной скоростью v

(рис. 8), получает приказ подняться на поверхность. Если промежуток времени, за который цистерны подлодки освобождаются от воды и заполняются воздухом, с тем чтобы ее средняя плотность ρ_1 стала меньше плотности воды ρ_0, невелик, то можно считать, что в момент $t = 0$ на подлодку начинает действовать выталкивающая сила, большая, чем вес лодки. По закону Архимеда выталкивающая сила равна $F = gV\rho_0$, где g — ускорение свободного падения, V — объем подлодки. Суммарная сила, действующая на подлодку в вертикальном направлении, — разность между F и весом тела $P = gV\rho_1$, а сообщаемое ею ускорение по второму закону Ньютона равно

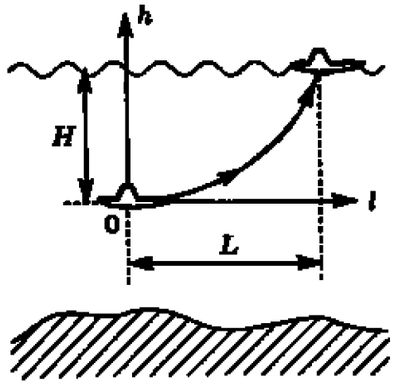


Рис. 8

$$\rho_1 V\frac{d^2h}{dt^2} = F - P = gV(\rho_0 - \rho_1).$$

Координата l, характеризующая горизонтальное положение подлодки, изменяется по закону движения тела с постоянной скоростью:

$$\frac{dl}{dt} = v.$$

Решая эти уравнения, находим, что

$$h(t) = g\,\frac{\rho_0 - \rho_1}{\rho_1}\,t^2, \qquad l(t) = vt, \tag{1}$$

и что лодка всплывет на поверхность в момент $t = t_k$, когда

$$h(t_k) = g\,\frac{\rho_0 - \rho_1}{\rho_1}\,t_k^2 = H, \qquad t_k = \left(\frac{\rho_1 H}{g(\rho_0 - \rho_1)}\right)^{1/2}.$$

При этом в горизонтальном направлении подлодка пройдет расстояние

$$L = vt_k = \left(\frac{\rho_1 H}{g(\rho_0 - \rho_1)}\right)^{1/2}.$$

Исключая из (1) время, найдем траекторию движения подлодки в координатах (l, h)

$$h = g\,\frac{\rho_0 - \rho_1}{\rho_1 v^2}\,l^2,$$

которая оказывается параболой с вершиной в точке $l = 0$, $h = 0$ (при выводе (1) вертикальная скорость лодки, а также величины l и h принимались равными нулю в момент $t = 0$). Считалось также, что никакие другие вертикальные силы, кроме F и P, на подлодку не действуют. Это предположение верно лишь при малых скоростях всплытия, когда можно пренебречь сопротивлением воды движению лодки (см. упр. 1).

Итак, непосредственное применение закона Архимеда, определяющего величину выталкивающей силы, и закона Ньютона, связывающе-

го силу, действующую на тело, и его ускорение, позволило легко найти траекторию подлодки Очевидно, что параболической траекторией обладает любое движущееся в плоскости тело, имеющее по одному из направлений постоянную скорость и на которое в другом направлении действует постоянная сила (уравнения (1) фактически дают параметрическую запись параболы) К таким движениям относятся, например, полет камня, брошенного с высоты H с горизонтальной скоростью v или полет электрона в электрическом поле плоского конденсатора Однако в последнем случае получить траекторию тела непосредственно из фундаментальных законов нельзя, требуется применить более детальную процедуру Рассмотрим этот вопрос подробнее

2. Отклонение заряженной частицы в электронно-лучевой трубке. Будем считать, что обкладки конденсатора электронно-лучевой трубки (рис 9) представляют собой бесконечные плоскости

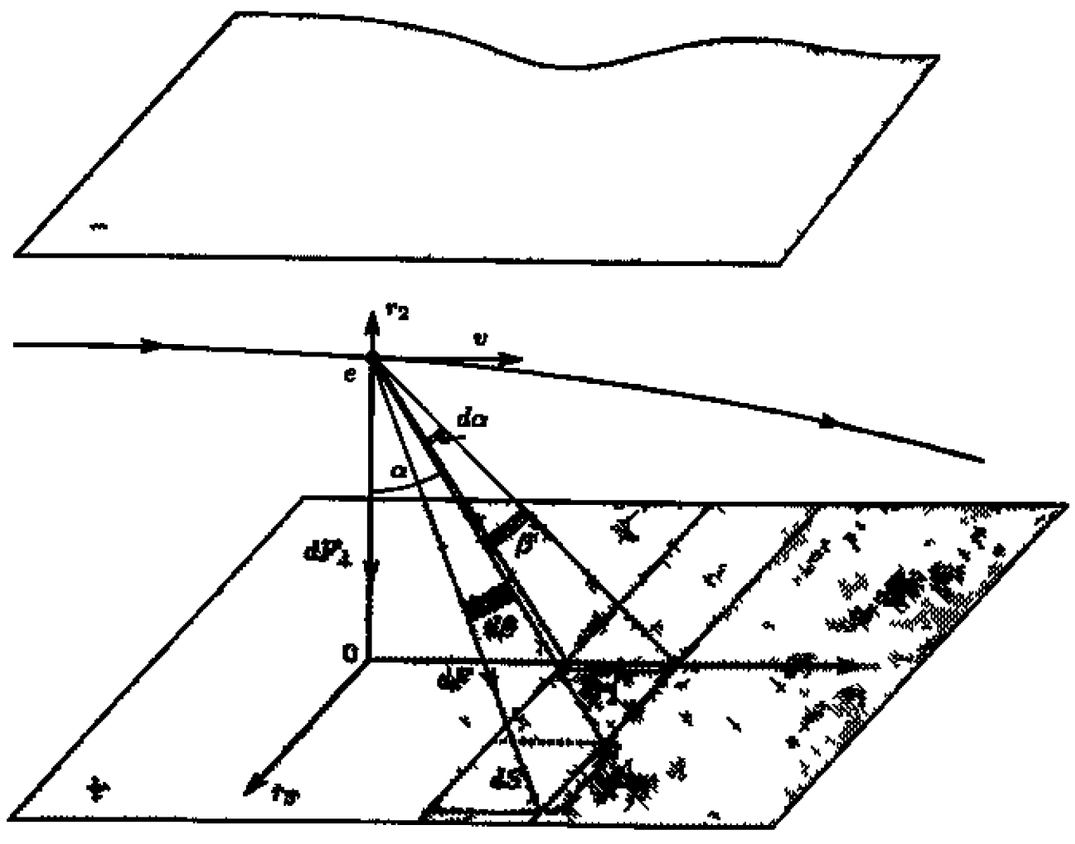


Рис 9

(предположение справедливо в случае, если расстояние между обкладками много меньше их размеров, а электрон движется на большом удалении от их краев, упр 2) Очевидно, что электрон будет притягиваться к нижней обкладке и отталкиваться от верхней Сила притяжения F двух разноименных зарядов элементарно определяется из закона Кулона

$$F = \frac{q_1 q_2}{r^2},$$

где q_1 и q_2 — величины зарядов, r — расстояние между ними Сложность заключается в том, что в данном примере на обкладке находится бесконечно много зарядов, каждый из которых расположен на своем расстоянии от движущегося электрона Поэтому необходимо сначала найти силу, индуцируемую каждым зарядом, и затем, просуммировав все элементарные силы, определить результирующее действие обкладок на электрон

Разобъем всю плоскость нижней обкладки на элементарные «полоски», характеризующиеся координатами r_1, r_2, r_3, $-\infty < r_1, r_3 < \infty$, $r_2 \equiv 0$ (см рис 9)

Подсчитаем силу притяжения электрона зарядом, находящемся на элементарной площадке $ds = dr_1\, dr_3$ и равным $dq = q_0\, ds$, где q_0 — поверхностная плотность заряда на обкладке Если частица находится на расстоянии r_2 от заряженной плоскости, то

$$dr_1 = r_2\,(\operatorname{tg}(\alpha + d\alpha) - \operatorname{tg}\alpha) = r_2\,\frac{d\alpha}{\cos^2\alpha}$$

(здесь учитывается малость величины $d\alpha$) Для определения величины dr_3 имеем

$$\frac{r_3 + dr_3}{r_1 + dr_1} = \frac{\operatorname{tg}(\beta + d\beta)}{\sin(\alpha + d\alpha)}, \qquad \frac{r_3}{r_2} = \frac{\operatorname{tg}\beta}{\operatorname{tg}\alpha}$$

Из последних двух формул находим

$$dr_3 = (r_1 + dr_1)\operatorname{tg}(\beta + d\beta) - r_1 \operatorname{tg}\beta = \frac{r_1\, d\beta/(\cos^2\beta) + dr_1 \operatorname{tg}\beta}{\sin\alpha},$$

где, аналогично предыдущему, учтена также и малость величины $d\beta$ Умножая dr_1 на dr_3 и отбрасывая член более высокого порядка малости, получаем $ds = r_2 r_1\, d\alpha\, d\beta/(\cos^2\alpha\cos^2\beta\sin\alpha)$ Сила притяжения электрона с зарядом q_e к элементарной площадке ds равна

$$dF = \frac{q_e q_0 r_2 r_1\, d\alpha\, d\beta}{\bar{r}^2\cos^2\alpha\cos^2\beta\sin\alpha},$$

где $\bar{r}$ — «среднее» расстояние от электрона до площадки, которое с учетом малости величин $d\alpha$, $d\beta$ вычисляется по формуле $\bar{r} = r_2/(\cos\alpha\cos\beta)$ В итоге для элементарной силы имеем

$$dF = q_e q_0\,\frac{r_1}{r_2}\,\frac{d\alpha\, d\beta}{\sin\alpha} = \frac{q_e q_0}{\cos\alpha}\, d\alpha\, d\beta,$$

а для ее вертикальной составляющей

$$dF_\perp = dF\cos\beta\cos\alpha = q_e q_0 \cos\beta\, d\alpha\, d\beta$$

Проинтегрировав выражение для $F_\perp$ по β от $\beta = 0$ до $\beta = \pi/2$, найдем силу притяжения электрона к части элементарной «полоски», расположенной в квадранте $r_1 > 0$, $r_3 > 0$

$$dF_\alpha^+ = q_e q_0\, d\alpha$$

Просуммировав $dF^+{}_\alpha$ по α от $\alpha = 0$ до $\alpha = \pi/2$, т. е. по всем полоскам квадранта $r_1 > 0$, $r_3 > 0$, определим силу притяжения, индуцируемую зарядами, расположенными в этом квадранте:

$$dF^+ = \frac{\pi}{2} q_e q_0.$$

Учитывая действие всех четырех квадрантов плоскости нижней обкладки и проводя аналогичные рассуждения для верхней обкладки, получим результирующую силу притяжения (отталкивания) электрона ко всем зарядам конденсатора

$$F = 4\pi q_e q_0. \qquad (2)$$

Сила F направлена вдоль оси r_2 к нижней обкладке (составляющие F по осям r_1, r_3, очевидно, равны нулю в силу симметрии — чтобы убедиться в этом, достаточно рассмотреть действие заряда, находящегося на площадке, расположенной в квадранте $r_1 < 0$, $r_3 < 0$ и симметричной площадке ds).

Поскольку сила F не зависит от r_2, а по горизонтальной оси частица движется с постоянной скоростью v, то приходим к ситуации предыдущего пункта — применив второй закон Ньютона, легко получить формулы, аналогичные (1), описывающие движение электрона по параболической траектории и дающие возможность вычислить все ее параметры. Однако в отличие от случая с подлодкой прямое применение фундаментального закона Кулона для получения модели движения электрона оказывается невозможным. Потребовалось, опираясь на фундаментальный закон, сначала описать элементарный акт взаимодействия зарядов, и уж затем, просуммировав все эти акты, удалось найти результирующую силу.

Подобная ситуация и последовательность действий весьма типичны при построении моделей, так как многие фундаментальные законы устанавливают взаимоотношения как раз между элементарными частями исходного объекта. Это, разумеется, справедливо не только для электрических сил, но, например, и для сил тяготения.

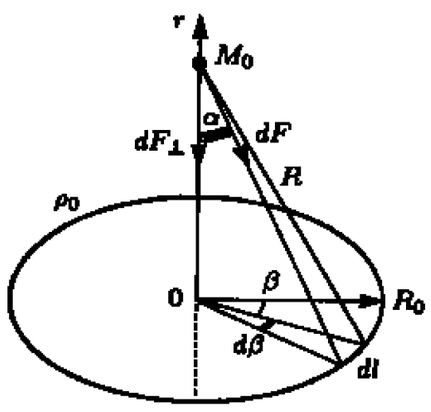


Рис. 10

3. Колебания колец Сатурна. Построим модель движения точечной массы M_0 в поле сил тяготения, создаваемом материальным кольцом с радиусом R_0 и линейной плотностью ρ_0. Кольцо считается бесконечно тонким, движение происходит вдоль оси кольца (рис. 10). Данная схема может рассматриваться как идеализация процесса колебаний колец Сатурна. Тем не менее, несмотря на существенные упрощения, непосредственное использование закона всемирного тяготения

$$F = \gamma \frac{m_0 m_1}{r^2},$$

где F — сила притяжения двух тел, имеющих массы m_0 и m_1, r — расстояние между ними, γ — постоянная тяготения, не может дать окончательной модели движения колец Сатурна, так как массы m_0, m_1 должны быть точечными.

Поэтому вычислим сначала силу притяжения между точечной массой M_0 и массой dm, содержащейся в малом элементе кольца dl, которую уже можно считать точечной:

$$dF = \gamma \frac{M_0\, dm}{R^2}.$$

Здесь R, r — соответственно расстояние от массы M_0 до кольца и до центра кольца. Очевидно, что при $0 \leqslant \alpha \leqslant \pi/2$ (для $\pi/2 \leqslant \alpha \leqslant 2\pi$ выкладки аналогичны)

$$\frac{R_0}{R} = \sin\alpha = \frac{R_0}{\sqrt{r^2 + R_0^2}}, \qquad \frac{r}{R} = -\cos\alpha = \frac{r}{\sqrt{r^2 + R_0^2}}.$$

Поскольку $dm = \rho_0\, dl = \rho_0 R_0\, d\beta = -\rho_0 r \operatorname{tg}\alpha\, d\beta$, то

$$dF = -\gamma \frac{M_0\rho_0}{R^2} r \operatorname{tg}\alpha\, d\beta = -\gamma \frac{M_0\rho_0}{r} \sin\alpha\cos\alpha\, d\beta.$$

Найдем проекцию силы dF на ось r (именно эта проекция определяет интересующее нас движение):

$$dF_\perp = dF\cos\alpha = -\gamma \frac{M_0\rho_0}{r} \sin\alpha\cos^2\alpha\, d\beta.$$

Просуммировав теперь силы тяготения, создаваемые всеми элементами кольца, т. е. взяв интеграл от $dF_\perp$ по β от $\beta = 0$ до $\beta = 2\pi$, найдем результирующую силу:

$$F = -2\pi\gamma \frac{M_0\rho_0}{r} \sin\alpha\cos^2\alpha = -\gamma M_0 M_1 \frac{r}{(r^2 + R_0^2)^{3/2}}, \qquad (3)$$

где $M_1 = 2\pi R_0\rho_0$ — полная масса кольца. Как и в предыдущем пункте, горизонтальная проекция результирующей силы равна нулю из-за симметричного расположения кольца относительно массы M_0.

Сила тяготения (3) существенно отличается от выражения, даваемого законом для точечных масс, переходя в него лишь при $r \gg R_0$, когда кольцо можно уподобить точечной массе благодаря большому, в сравнении с размерами кольца, расстоянию между тяготеющими телами. Если же $r \ll R_0$, то

$$F = -\gamma \frac{M_0 M_1}{R_0^3} r,$$

и сила притяжения, в противоположность случаю точечных масс, убывает с уменьшением расстояния между объектами (еще один предельный переход рассмотрен в упр. 3).

Применив к массе M_0 второй закон Ньютона, получим уравнение ее движения вдоль оси r:

$$\frac{d^2r}{dt^2} = -\gamma M_1 \frac{r}{(r^2 + R_0^2)^{3/2}},$$

которое, в отличие от п. 1 и п. 2, существенно нелинейно и становится линейным лишь при $r \ll R_0$:

$$\frac{d^2r}{dt^2} = -\gamma \frac{M_1}{R_0^3} r. \qquad (4)$$

4. Движение шарика, присоединенного к пружине. В получении моделей пп. 1–3 главную роль играли фундаментальные законы, определявшие происхождение и величину сил, действующих на объект, а второй закон Ньютона был как бы вспомогательным и применялся на последней стадии построения модели. Конечно же, такое деление чисто условно. Ведь если речь идет о задачах динамики, то можно использовать и другую схему — сначала связать с помощью закона Ньютона проекции ускорения тела с проекциями действующих на него сил, а затем, исходя из тех или иных соображений, вычислить эти силы как функции координат, получив замкнутую модель. Продемонстрируем этот подход на примере модели движения шарика, присоединенного к пружине, с жестко закрепленным концом (рис. 11).

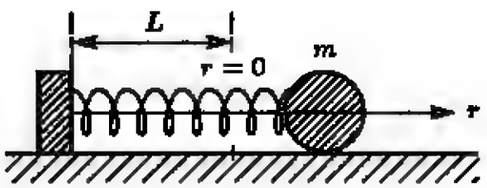


Рис. 11

Пусть r — координата шарика вдоль оси пружины, лежащей на горизонтальной плоскости, и направление движения шарика совпадает с ее осью. Тогда по второму закону динамики

$$F = ma = m\frac{d^2r}{dt^2},$$

где m — масса шарика, a — его ускорение. Будем считать плоскость идеально гладкой (т. е. движение происходит без трения), пренебрежем также сопротивлением воздуха и примем во внимание то, что вес шарика уравновешивается реакцией плоскости. Единственная сила, действующая на шарик в направлении оси r, очевидно, сила упругости пружины. Определим ее, используя закон Гука, гласящий, что для растяжения (сжатия) пружины необходимо приложить силу

$$F = -kr,$$

где коэффициент $k > 0$ характеризует упругие свойства пружины, а r — величну ее растяжения или сжатия относительно нейтрального, ненагруженного положения $r = 0$. Уравнение движения шарика принимает вид (уравнение элементарного осциллятора)

$$m\frac{d^2r}{dt^2} = -kr, \qquad t > 0. \qquad (5)$$

Оно описывает его гармонические колебания и имеет общее решение

$$r = A\sin\omega t + B\cos\omega t, \qquad (6)$$

где $\omega = \sqrt{k/m}$ — собственная частота колебаний системы «пружина—шарик». Значения A и B легко определяются из начального состояния объекта, т. е. через величины $r(t=0) = r_0$ и $v(t=0) = v_0$ ($v(t)$ — скорость шарика), причем $r(t) \equiv 0$ при $r_0 = v_0 = 0$. Заметим, что уравнение (4) с точностью до обозначений совпадает с (5), поэтому в п. 3 речь также шла о процессе колебаний, но применительно к системе «Сатурн—кольцо».

Подходы, с помощью которых строились модели данного параграфа, не должны, разумеется, противоречить другим фундаментальным законам природы. Соответствующая проверка непротиворечивости (если она возможна) весьма полезна для установления правильности моделей. Поясним это, используя для вывода уравнения (5) не закон Ньютона, а закон сохранения энергии. Поскольку точка крепления пружины неподвижна, то стенка не совершает работу над системой «пружина—шарик» (и наоборот), и ее полная механическая энергия E остается постоянной. Вычислим ее. Кинетическая энергия определяется движением шарика (пружина считается невесомой):

$$E_к = \frac{mv^2}{2} = \frac{m(dr/dt)^2}{2}.$$

Потенциальная энергия системы «содержится» в пружине, ее нетрудно найти, определив работу, необходимую для растяжения (сжатия) пружины на величину r:

$$E_п = -\int_0^r F dr' = \int_0^r kr'\,dr' = k\frac{r^2}{2}.$$

Для неизменной со временем величины $E = E_к + E_п$ (интеграла энергии) получаем

$$E = \frac{m(dr/dt)^2}{2} + \frac{kr^2}{2}.$$

Так как $dE/dt \equiv 0$, то, продифференцировав интеграл энергии по t, приходим к выражению

$$m\frac{dr}{dt}\frac{d^2r}{dt^2} + k\frac{dr}{dt}r = \frac{dr}{dt}\left(m\frac{d^2r}{dt^2} + kr\right) = 0,$$

т. е. к уравнению (5), проверив тем самым правильность его получения. Подобную процедуру нетрудно провести и для примеров пп. 1–3.

5. Заключение.

1. Даже в простейших ситуациях для построения модели может потребоваться использование не одного, а нескольких фундаментальных законов.

2. Прямое формальное применение фундаментальных законов к объекту, рассматриваемому как целое, не всегда возможно (пп. 2, 3). В этих случаях требуется просуммировать элементарные акты взаимодействия между его частями, принимая во внимание свойства объекта (например, его геометрию).

3. Одними и теми же моделями могут описываться совершенно разные по своей природе объекты, подчиняющиеся разным фундаментальным законам, и, наоборот, данному закону могут отвечать принципиально разные модели (например, линейные и нелинейные; см. п. 3).

4. Необходимо использовать все возможности для проверки правильности построения модели (предельные переходы — пп. 2, 3, другие фундаментальные законы — п. 4 и т. д.).

УПРАЖНЕНИЯ

1. В задаче о всплытии подводной лодки учитывается сопротивление воды. Принимая силу сопротивления равной $F_1 = -k_0 u$, где $k_0 > 0$ — коэффициент, зависящий от свойств воды и формы подлодки, u — вертикальная скорость лодки, найдите максимальную глубину H, при всплытии с которой силой F_1 можно пренебречь в любой момент времени $t \leqslant t_k$ (должно выполняться требование $F_1 \ll F - P$).

2. Повторяя рассуждения п. 2, найдите силу притяжения электрона к обкладкам конденсатора, имеющим конечные размеры R_1, R_3. Убедитесь в том, что при $R_1 \to \infty$, $R_3 \to \infty$ полученное выражение переходит в формулу (2)

3. В задаче п. 3 введите толщину кольца d, найдите силу F и убедитесь, что полученное выражение при $d \to 0$ совпадает с формулой (3).

4. Пусть расстояние между точкой нейтрального положения пружины $r = 0$ и стенкой, к которой она крепится, равно L (см. рис. 11). Найдите, пользуясь формулой (6), условия на величины r_0, v_0, при выполнении которых шарик не может удариться о стенку (в противном случае модель (5) неверна, так как при соударении со стенкой шарик испытывает с ее стороны действие некоторой силы, не учитываемой в уравнении (5)).

§ 3. Вариационные принципы и математические модели

Дадим упрощенную формулировку вариационного принципа Гамильтона для механической системы. На его основе выведем уравнения движения шарика на пружине и маятника в поле сил тяжести. Сопоставим результаты получения моделей из фундаментальных законов и из вариационного принципа.

1. Общая схема принципа Гамильтона. Пусть имеется механическая система, формального и строго определения которой пока давать не будем, имея в виду, однако, что все элементы такой системы и взаимодействие между ними определяются законами механики (один из простейших примеров — рассмотренная в п. 4 § 2 система «шарик—пружина»). Введем понятие *обобщенных координат* $Q(t)$, полностью определяющих положение механической системы в пространстве. Величина $Q(t)$ может быть декартовой координатой (например, координата r в системе «шарик—пружина»), радиусом-вектором, угловой координатой, набором координат материальных точек, составляющих

систему, и т. д. Величину dQ/dt естественно назвать *обобщенной скоростью* механической системы в момент времени t. Набор величин $Q(t)$ и dQ/dt определяет состояние механической системы во все моменты времени.

Для описания механической системы вводится *функция Лагранжа*, построение которой — отдельный вопрос, более подробно рассматриваемый в гл. III. В простейших случаях функция Лагранжа имеет ясный смысл и записывается в виде

$$L(Q, dQ/dt) = E_{\text{к}} - E_{\text{п}}, \tag{1}$$

где $E_{\text{к}}$, $E_{\text{п}}$ — кинетическая и потенциальная энергии системы соответственно. Для целей данного параграфа нет необходимости давать общее определение величин $E_{\text{к}}$, $E_{\text{п}}$, поскольку в рассматриваемых примерах они вычисляются очевидным образом.

Введем далее величину $S[Q]$, называемую *действием*:

$$S[Q] = \int_{t_1}^{t_2} L\left(Q, \frac{dQ}{dt}\right) dt \tag{2}$$

Интеграл (2), очевидно, является функционалом от обобщенной координаты $Q(t)$, т. е. функции $Q(t)$, заданной на отрезке $[t_1, t_2]$, он ставит в соответствие некоторое число S (действие).

Принцип Гамильтона для механической системы гласит: если система движется по законам механики, то $Q(t)$ — стационарная функция для $S[Q]$, или

$$\frac{d}{d\varepsilon} S[Q + \varepsilon\varphi]_{\varepsilon=0} = 0. \tag{3}$$

Фигурирующая в *принципе наименьшего действия* (3) функция $\varphi(t)$ — некоторая пробная функция, обращающаяся в нуль в моменты t_1, t_2 и удовлетворяющая тому условию, что $Q(t) + \varepsilon\varphi(t)$ — возможная координата данной системы (в остальном $\varphi(t)$ произвольна).

Смысл принципа (3) в том, что из всех априори мыслимых (допускаемых) траекторий (движений) системы между моментами t_1, t_2 выбирается (реализуется) движение, доставляющее минимум функционалу действия (отсюда происходит и название принципа). Функция $\varepsilon\varphi(t)$ называется *вариацией* величины $Q(t)$.

Итак, схема применения принципа Гамильтона (3) для построения моделей механических систем состоит в следующем: определяются обобщенные координаты $Q(t)$ и обобщенные скорости dQ/dt системы, строятся функция Лагранжа $L(Q, dQ/dt)$ и функционал действия $S[Q]$, минимизация которого на вариациях $\varepsilon\varphi(t)$ координаты $Q(t)$ и дает искомую модель.

2. Третий способ получения модели системы «шарик—пружина». Воспользуемся принципом Гамильтона для построения модели движения шарика, соединенного с пружиной (п. 4 § 2). В качестве обобщенной координаты системы естественно выбрать обычную

эйлерову координату шарика $r(t)$. Тогда обобщенная скорость $dr/dt = v(t)$ — обычная скорость шарика. Функция Лагранжа (1), равная $L = E_к - E_п$, записывается через уже найденные в п. 4 § 2 значения кинетической и потенциальной энергии системы:

$$L = \frac{m(dr/dt)^2}{2} - k\frac{r^2}{2}.$$

Для величины действия получаем выражение

$$S[r] = \int_{t_1}^{t_2} L\left(r, \frac{dr}{dt}\right) dt = \int_{t_1}^{t_2} \left[\frac{m}{2}\left(\frac{dr}{dt}\right)^2 - \frac{k}{2}r^2\right] dt.$$

Теперь, в соответствии со схемой п. 1, вычислим действие на вариациях $\varepsilon\varphi(t)$ координаты $r(t)$:

$$S[r+\varepsilon\varphi] = \int_{t_1}^{t_2} \left[\frac{m}{2}\left(\frac{d(r+\varepsilon\varphi)}{dt}\right)^2 - \frac{k}{2}(r+\varepsilon\varphi)^2\right] dt.$$

Последнюю формулу необходимо продифференцировать по ε (учитывая, что функции r, φ, dr/dt, $d\varphi/dt$ от ε не зависят):

$$\frac{d}{d\varepsilon} S[r+\varepsilon\varphi] = \frac{d}{d\varepsilon}\frac{1}{2}\int_{t_1}^{t_2} \left[m\left\{\left(\frac{dr}{dt}\right)^2 + 2\varepsilon\frac{dr}{dt}\frac{d\varphi}{dt} + \varepsilon^2\left(\frac{d\varphi}{dt}\right)^2\right\} - \right.$$

$$\left. - k\left\{r^2 + 2\varepsilon r\varphi + \varepsilon^2\varphi^2\right\}\right] dt =$$

$$= \int_{t_1}^{t_2} \left[m\left\{\frac{dr}{dt}\frac{d\varphi}{dt} + \varepsilon\left(\frac{d\varphi}{dt}\right)^2\right\} - k\left\{r\varphi + \varepsilon\varphi^2\right\}\right] dt,$$

и положить в ней $\varepsilon = 0$:

$$\frac{d}{d\varepsilon} S[r+\varepsilon\varphi]\Big|_{\varepsilon=0} = \int_{t_1}^{t_2} \left[m\frac{dr}{dt}\frac{d\varphi}{dt} - kr\varphi\right] dt = 0.$$

Правая часть этого выражения (равного нулю в согласии с принципом Гамильтона — см. (3)) с помощью интегрирования ее первого члена по частям и с учетом того, что $\varphi = 0$ в моменты t_1, t_2, преобразуется к виду

$$\frac{d}{d\varepsilon} S[r+\varepsilon\varphi]\Big|_{\varepsilon=0} = -\int_{t_1}^{t_2} \varphi\left[m\frac{d^2r}{dt^2} + kr\right] dt = 0.$$

Поскольку пробная функция $\varphi(t)$, фигурирующая в формулировке принципа наименьшего действия, произвольна, то часть выражения, стоящая под знаком интеграла в квадратных скобках, должна быть равна нулю во все моменты времени $t_1 < t < t_2$:

$$m\frac{d^2r}{dt^2} = -kr,$$

т. е. движение системы должно описываться уравнением (5), полученным в § 2 из закона Ньютона (первый способ) и закона сохранения энергии (второй способ). Все три подхода оказываются эквивалентными (так как между ними существует глубокая связь, более подробно изучаемая в гл. III).

3. Колебания маятника в поле сил тяжести. Приведем несколько более сложный пример применения принципа Гамильтона с подробным рассмотрением начальной стадии построения модели — описанием механической системы.

Пусть на неподвижном шарнире подвешен маятник — груз массы m, находящийся на конце стержня длины l (рис. 12). Шарнир считается идеально гладким в том смысле, что в нем не происходят потери энергии на трение. Неподвижность шарнира означает, что от него энергия в систему «стержень—груз» не поступает, такой шарнир неспособен совершить над ней какую-либо работу. Стержень считается невесомым и абсолютно жестким, т. е. его кинетическая и потенциальная энергии равны нулю, а груз не может совершать движений вдоль оси стержня. Груз имеет небольшие размеры по сравнению с длиной стержня (материальная точка), ускорение свободного падения g постоянно, сопротивлением воздуха пренебрегается, колебания происходят в фиксированной вертикальной плоскости (для чего, очевидно, вектор начальной скорости груза должен лежать в этой плоскости).

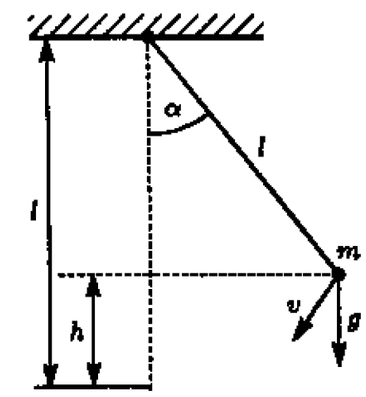


Рис. 12

После всех этих упрощающих предположений ясно, что положение маятника определяется лишь одной обобщенной координатой, в качестве которой выберем угол $\alpha(t)$ отклонения стержня от вертикали. Обобщенная скорость в данном случае — угловая скорость $d\alpha/dt$.

Кинетическая энергия системы дается формулой

$$E_к = \frac{1}{2}mv^2 = \frac{1}{2}m\left(l\frac{d\alpha}{dt}\right)^2 = \frac{1}{2}ml^2\left(\frac{d\alpha}{dt}\right)^2,$$

а потенциальная энергия выражением

$$E_п = mgh = -mg(l\cos\alpha - l),$$

где h — отклонение маятника от наинизшего положения по вертикали. В дальнейших выкладках величину mgl в $E_п$ опустим, так как потенциальная энергия определяется с точностью до постоянной.

Теперь нетрудно вычислить функцию Лагранжа (1) и действие (2):

$$L\left(\alpha, \frac{d\alpha}{dt}\right) = ml\left[\frac{1}{2}l\left(\frac{d\alpha}{dt}\right)^2 + g\cos\alpha\right],$$

$$S[\alpha] = ml\int\limits_{t_1}^{t_2}\left[\frac{1}{2}l\left(\frac{d\alpha}{dt}\right)^2 + g\cos\alpha\right]dt.$$

Находя действие на вариациях $\alpha + \varepsilon\varphi(t)$:

$$S[\alpha+\varepsilon\varphi] = ml\int\limits_{t_1}^{t_2}\left[\frac{1}{2}l\left(\frac{d\alpha}{dt} + \varepsilon\frac{d\varphi}{dt}\right)^2 + g\cos(\alpha+\varepsilon\varphi)\right]dt =$$

$$= ml\int\limits_{t_1}^{t_2}\left[\frac{1}{2}\left\{\left(\frac{d\alpha}{dt}\right)^2 + 2\varepsilon\frac{d\alpha}{dt}\frac{d\varphi}{dt} + \varepsilon^2\left(\frac{d\varphi}{dt}\right)^2\right\} + g\cos(\alpha+\varepsilon\varphi)\right]dt,$$

дифференцируя его по ε и полагая $\varepsilon = 0$, получаем

$$\frac{d}{d\varepsilon}S[\alpha+\varepsilon\varphi]\Big|_{\varepsilon=0} = ml\int\limits_{t_1}^{t_2}\left[l\frac{d\alpha}{dt}\frac{d\varphi}{dt} - \varphi g\sin\alpha\right]dt = 0.$$

Как и в п. 1, интегрируем первый член выражения в скобках по частям и, учитывая, что $\varphi(t) = 0$ в моменты t_1, t_2, приходим к следующему уравнению:

$$ml\int\limits_{t_1}^{t_2}\varphi\left[l\frac{d^2\alpha}{dt^2} + g\sin\alpha\right]dt = 0,$$

которое в силу произвольности $\varphi(t)$ может удовлетворяться лишь если для всех $t_1 < t < t_2$ справедливо

$$\frac{d^2\alpha}{dt^2} = -\frac{g}{l}\sin\alpha. \tag{4}$$

Заметим, что уравнение колебаний маятника (4) в отличие от уравнения (5) § 2 нелинейно. Это обстоятельство связано с более сложной геометрией системы «стержень—груз», а именно: ускорение, испытываемое грузом, не пропорционально координате, как в случае закона Гука, а является более сложной функцией отклонения от положения равновесия (угла α). Если же эти отклонения малы, то $\sin\alpha \approx \alpha$, и модель малых колебаний линейна:

$$\frac{d^2\alpha}{dt^2} = -\frac{g}{l}\alpha.$$

Они описываются формулой, аналогичной (6) из § 2, где $\omega = \sqrt{g/l}$ — собственная частота малых колебаний, а величины A, B определяются через $\alpha(t=0)$, $\frac{d\alpha}{dt}(t=0)$.

4. Заключение. Примеры использования принципа Гамильтона для построения моделей механических систем рисуют весьма четкую программу действий, в общем виде описанную в п. 1. Универсальность, строго формализованные последовательные процедуры, не зависящие от деталей конкретной системы, безусловно, весьма привлекательная черта вариационных принципов. В приведенных выше простых случаях модели могут быть относительно легко получены и иными способами. Однако для многих других, более сложных объектов, вариационные принципы оказываются фактически единственным методом построения моделей. Так, например, механические части большинства робототехнических устройств состоят из большого количества разнообразных элементов, связанных между собой различными способами. Их математические модели включают большое число уравнений, единообразно получаемых в основном с помощью вариационных принципов. Этот подход успешно применяется также и для систем иной природы (физических, химических, биологических), для которых формулируются соответствующие общие утверждения о характере их эволюции (поведения).

То обстоятельство, что принцип Гамильтона и другие подходы дают совпадающие модели, естественно, поскольку они описывают один и тот же исходный объект. Разумеется, такое совпадение гарантировано только при одних и тех же исходных предположениях об объекте. Если его идеализация (как один из первых этапов построения модели) проводится одинаково, то разные способы получения моделей должны давать тождественные результаты. Пусть, например, в системе «шарик—пружина» появляется дополнительная сила некоторого внешнего воздействия на шарик F_1. Тогда из второго закона Ньютона нетрудно получить уравнение движения шарика

$$m\frac{d^2r}{dt^2} = -kr + F_1$$

(ср. с (5) § 2). Применяя принцип Гамильтона к такой системе, необходимо учесть наличие этой силы. Очевидно, что определения обобщенной координаты, обобщенной скорости и кинетической энергии $E_к$ остаются неизменными. В то же время выражение для потенциальной энергии существенно изменяется (ср. с п. 2) на величину, равную работе, произведенной этой силой над системой:

$$E_п = k\frac{r^2}{2} + \int\limits_0^r F_1\,dr = k\frac{r^2}{2} + F_1 r.$$

Проводя аналогичные п. 2 выкладки с соответствующим образом измененными величинами L и Q, нетрудно убедиться в том, что принцип Гамильтона дает написанное выше уравнение с внешней силой F_1.

УПРАЖНЕНИЯ

1. Проверьте, что модель (4) построена правильно, получив ее с помощью второго закона Ньютона.

2. Используя результаты п. 2 § 2 и принцип Гамильтона, постройте модель колебаний маятника в электрическом поле, создаваемом заряженной горизонтальной плоскостью, над которой подвешен маятник. Заряд груза равен q, поверхностная плотность зарядов на плоскости — q_0 (силой тяжести пренебречь). Почему, несмотря на различную природу действующих сил, получается модель, аналогичная (4)?

§ 4. Пример иерархии моделей

Для движения шарика, соединенного с пружиной, построим иерархическую цепочку моделей по принципу «снизу вверх». Последовательно введем новые усложняющие факторы и дадим их математическое описание.

1. Различные варианты действия заданной внешней силы. Пусть на шарик действует известная внешняя сила $F(r,t)$, зависящая от времени и положения шарика. Она может порождаться полем тяготения (см. упр. 1), иметь электрическое или магнитное происхождение и т. д. Из второго закона Ньютона сразу получаем, что по сравнению с базовой моделью колебаний

$$m\frac{d^2r}{dt^2} = -kr \tag{1}$$

в правой части уравнения (1) появляется дополнительный член:

$$m\frac{d^2r}{dt^2} = -kr + F(r,t). \tag{2}$$

Простейший вариант уравнения (2) отвечает постоянной силе $F(r,t) = F_0$. Проведя замену $\bar{r} = r - F_0/k$, получаем для $\bar{r}$

$$\frac{d^2\bar{r}}{dt^2} = -k\bar{r},$$

т. е. постоянная сила не вносит изменений в процесс колебаний за тем исключением, что координата нейтральной точки, в которой сила, действующая на шарик, равна нулю, сдвигается на величину F_0/k.

Гораздо более сложная картина движения может порождаться зависящей от времени силой $F(t)$. Рассмотрим для определенности периодическую внешнюю силу $F(t) = F_0 \sin\omega_1 t$:

$$m\frac{d^2r}{dt^2} = -kr + F(t) = -kr + F_0\sin\omega_1 t. \tag{3}$$

Решение линейного уравнения (3) находится как сумма общего решения однородного уравнения (формула (6) § 2) и частного решения неоднородного уравнения (3), которое будем искать в виде

$$r_1(t) = C\sin\omega_1 t. \tag{4}$$

Подстановкой этого выражения в (3) находим

$$C = \frac{F_0}{k - m\omega_1^2} = \frac{F_0}{m(\omega^2 - \omega_1^2)},$$

где $\omega = \sqrt{k/m}$ — частота колебаний пружины в отсутствие внешних сил, или собственная частота системы. В итоге для общего решения (3) имеем

$$r(t) = A\sin\omega t + B\cos\omega t + \frac{F_0}{m(\omega^2 - \omega_1^2)}\sin\omega_1 t.$$

Итак, внешняя сила $F(t)$ приводит не только к появлению в системе дополнительных колебаний с частотой ω_1, но и к возникновению резонанса — неограниченному росту амплитуды колебаний при $\omega_1 \to \omega$.

2. Движение точки крепления, пружина на вращающемся стержне. Резонанс в системе может быть вызван также благодаря действию сил инерционного происхождения. Пусть точка крепления пружины движется по заданному закону $r_0(t) = f(t)$. Тогда в системе координат, связанной с этой точкой, на шарик действует, помимо натяжения пружины, сила инерции, равная $ma(t)$, где $a(t)$ — ускорение, обусловленное движением системы координат, $a(t) = d^2f/dt^2$. В этой системе координат движение шарика описывается уравнением

$$m\frac{d^2r}{dt^2} = -kr + F(t),$$

где $F(t) = -ma(t) = -m\,d^2f/dt^2$ — некоторая заданная функция времени. Как и в предыдущем случае, при соответствующем периодическом движении точки крепления в системе возникает резонанс.

При более сложной геометрии силы инерции системы могут зависеть не только от времени t, но и от координаты r. Если пружина надета на стержень, движущийся с угловой скоростью $\omega(t)$, то центробежная сила инерции равна $F = mv^2(t)/R = m\omega^2(t)\,R$, где $v(t) = \omega(t)\,R$, $R = R_0 + r$, R_0 — длина пружины в ненагруженном состоянии, r — отклонение шарика от нейтрального положения, $r > -R_0$. Уравнение движения шарика принимает вид

$$m\frac{d^2r}{dt^2} = -kr + F(r,t), \tag{5}$$

где $F(r,t) = m\omega^2(t)\,(R_0 + r)$, или

$$m\frac{d^2r}{dt^2} = -(k - m\omega^2(t))\,r + m\omega^2(t)\,R_0,$$

причем, очевидно, при $r \ll R_0$ линейное уравнение (5), общее решение которого здесь не выписывается в силу его громоздкости, переходит в уравнение вида (3) с $F(t) = -m\omega^2(t)\,R_0$.

Однако в данном случае резонанс невозможен, так как внешняя сила всегда направлена в одну сторону и не в состоянии раскачать систему.

Заметим, что усложненная по сравнению с исходным случаем геометрия отнюдь не всегда означает более сложное поведение объекта. Рассмотрим, например, шарик, прикрепленный к двум пружинам с жесткостью k_1 и k_2 (рис. 13). Начало координат поместим в точку, где силы, действующие на шарик со стороны обеих пружин, уравновешивают друг друга (при этом должно соблюдаться некоторое условие на параметры системы, чтобы шарик не мог упираться в одну из точек крепления — см упр 2) По закону Гука при отклонении r на шарик со стороны левой пружины действует сила $-k_1 r$, а со стороны правой — сила $-k_2 r$ (обе силы направлены в одну сторону, так как при растяжении первой пружины вторая пружина, наоборот, сжимается). В итоге приходим к такому же уравнению, как и в случае одной пружины,

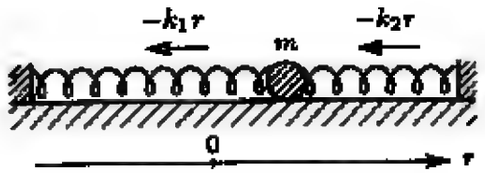


Рис 13

$$m\frac{d^2r}{dt^2} = -k_1 r - k_2 r = -kr,$$

но с увеличенной жесткостью $k = k_1 + k_2$, складывающейся из жесткостей обеих пружин

3. Учет сил трения. В рассматриваемой системе силы трения могут появляться по крайней мере из-за двух причин. Первая из них — неидеальность поверхностей шарика и плоскости, по которой он движется В этом случае сила трения равна $F = k_1 P$, где k_1 — коэффициент трения, $P = mg$ — вес шарика Она всегда направлена против движения шарика, ее знак противоположен знаку скорости шарика $v = dr/dt$, т е. $F = -k_1 mg\,\mathrm{sign}(dr/dt)$ Движение шарика подчиняется уравнению

$$m\frac{d^2r}{dt^2} = -kr - k_1 mg\,\mathrm{sign}\,\frac{dr}{dt}, \tag{6}$$

которое внешне похоже на уравнение (2) с постоянной силой $F(r,t) = F_0$ Однако из-за знакопеременности силы оно не сводится к стандартному уравнению колебаний. Это обстоятельство служит выражением того, что уравнения (1) и (6) описывают существенно разные процессы В частности, амплитуда колебаний шарика в последнем случае уменьшается со временем В этом нетрудно убедиться, переписав (6) в виде

$$m\frac{dv}{dt} + kr = -k_1 mg\,\mathrm{sign}\,v,$$

умножив обе части этого выражения на $v/2$.

$$m\frac{v}{2}\frac{dv}{dt} + kr\frac{v}{2} = -k_1 mg\,\mathrm{sign}\,v\,\frac{v}{2},$$

и с учетом того, что $v = dr/dt$, получив

$$\frac{m}{2}\frac{dv^2}{dt} + \frac{k}{2}\frac{dr^2}{dt} = -\frac{1}{2}k_1 mg\,\mathrm{sign}\,v\cdot v.$$

Последнее уравнение эквивалентно уравнению

$$\frac{d}{dt}\left(\frac{mv^2}{2} + k\frac{r^2}{2}\right) = -\frac{1}{2}k_1 mg\,\mathrm{sign}\,v\cdot v. \tag{7}$$

Принимая во внимание, что в левой части (7) под знаком производной стоит сумма кинетической и потенциальной энергии системы $E(t) = E_к(t) + E_п(t)$, а правая часть (7) при $v \neq 0$ отрицательна, имеем

$$\frac{dE(t)}{dt} < 0, \quad v \neq 0 \qquad \left(\frac{dE(t)}{dt} = 0, \quad v = 0\right),$$

т. е. полная энергия $E(t)$ убывает со временем. Поскольку в моменты достижения шариком максимальной амплитуды $r_m(t)$ его скорость (и кинетическая энергия $E_к$) равна нулю, то в эти моменты $E_п = kr_m^2(t)/2 = E(t)$, и в силу убывания $E(t)$ амплитуда $|r_m(t)|$ — также убывающая функция времени.

Более подробно рассмотрим результат действия силы трения иного происхождения, возникающей из-за сопротивления среды, в которой движется шарик (воздуха, воды и т. д.). В этом случае сила трения не постоянна, а существенно зависит от скорости движения. Эта зависимость описывается известной формулой Стокса

$$F = -\mu v = -\mu\frac{dr}{dt},$$

где коэффициент $\mu > 0$ определяется размерами шарика, плотностью среды, ее вязкостью и т. д. Уравнение движения в вязкой среде имеет вид

$$m\frac{d^2r}{dt^2} = -kr + F(v) = -kr - \mu\frac{dr}{dt}. \tag{8}$$

Найдем общее решение линейного уравнения (8), избавившись предварительно от члена с первой производной. Подстановка в (8) замены $r(t) = \bar{r}(t)\,e^{\alpha t}$ дает для новой функции $\bar{r}(t)$ уравнение

$$m\left(e^{\alpha t}\frac{d^2\bar{r}}{dt^2} + \alpha e^{\alpha t}\frac{d\bar{r}}{dt} + \alpha e^{\alpha t}\frac{d\bar{r}}{dt} + \alpha^2 e^{\alpha t}\bar{r}\right) =$$
$$= -k\bar{r}e^{\alpha t} - \mu e^{\alpha t}\frac{d\bar{r}}{dt} - \mu\alpha e^{\alpha t}\bar{r}.$$

Сократив в нем множитель $e^{\alpha t}$ и положив $\alpha = -\mu/(2m)$, придем к уравнению

$$m\frac{d^2\bar{r}}{dt^2} = -\left(k - \frac{\mu^2}{4m}\right)\bar{r} = -k_1\bar{r}. \tag{9}$$

В отличие от уравнения (1) первый множитель в правой части (9) может менять знак в зависимости от значений параметров k, μ, m системы, что с учетом связи $r(t) = e^{\alpha t}\bar{r}(t)$ приводит к существенно иному ее поведению относительно стандартного случая.

При малой вязкости, т. е. при $k - \mu^2/(4m) = k_1 > 0$ решение $\bar{r}(t)$ дается формулой (6) § 2, и для $r(t)$ имеем

$$r = \bar{r}e^{\alpha t} = e^{-t\mu/(2m)}\,(A\sin\omega t + B\cos\omega t)\,,$$

где $\omega = (k_1/m)^{1/2}$, а константы A, B находятся через r_0, v_0. В системе происходят затухающие со временем колебания (см. также упр. 3) с частотой ω.

Если $k_1 = 0$, то величина $d\bar{r}/dt$ постоянна, или, что то же самое, $\bar{r}(t) = ct + c_1$. Для $r(t)$ с учетом начальных данных получаем

$$r(t) = e^{-t\mu/(2m)}\,(ct + c_1) = e^{-t\mu/(2m)}\left(\left(v_0 + \frac{\mu r_0}{2m}\right)t + r_0\right).$$

В данном случае колебания отсутствуют благодаря подавляющему действию сил вязкого трения. Система может лишь один раз пройти точку $r = 0$, для чего необходимо и достаточно выполнения условий $v_0 < -\mu r_0/(2m)$, $r_0 > 0$ или $v_0 > -\mu r_0/(2m)$, $r_0 < 0$, т. е. начальная скорость шарика должна быть достаточно велика и направлена к точке $r = 0$. При этом, очевидно, скорость шарика $v(t) = dr/dt$ может менять знак лишь один раз.

Наконец, при большой вязкости действие силы трения настолько значительно, что для любых r_0, v_0 шарик «застревает» в среде, никогда не проходя точку $r = 0$, а лишь односторонне приближаясь к ней при $t \to \infty$. Действительно, при $k_1 < 0$ решение уравнения (9) (см. также упр. 4) знакопостоянно (предположение о противном сразу же приводит к противоречию с уравнением), следовательно, величина $r(t)$ также не меняет знак. Поведение функции $\bar{r}(t)$ при $t \to \infty$ можно понять из свойств первого интеграла уравнения (9)

$$m\left(\frac{d\bar{r}}{dt}\right)^2 = -k_1\bar{r}^2 + \text{const},$$

который нетрудно получить, умножая обе части (9) на $d\bar{r}/dt$ и интегрируя один раз по t. Предположения о том, что $\bar{r}(t) \to \infty$ или $\bar{r}(t) \to C_1$ при $t \to \infty$ противоречат последнему равенству. Остается единственный вариант $\bar{r}(t) \to 0$, $t \to \infty$, и, таким образом, $r(t) \to 0$, $t \to \infty$.

Итак, движение системы в вязкой среде отличается большим по отношению к идеальной ситуации разнообразием, причем во всех случаях оно происходит с затуханием.

4. Два типа нелинейных моделей системы «шарик—пружина». Формула Стокса справедлива, строго говоря, только для установившихся движений, когда действие постоянной внешней силы уравновешивается силой вязкого трения так, что в итоге тело перемещается с постоянной скоростью. Вполне возможны ситуации, при которых сила сопротивления вязкой среды при малых скоростях меньше, а при больших скоростях больше, чем вычисляемая по формуле Стокса; например, $F(v) = -\mu v\,|v|^\alpha$, где $\mu > 0$, $\alpha > -1$. Тогда искомая величина $r(t)$ определяется из уравнения

$$m\frac{d^2r}{dt^2} = -kr + F(v) = -kr - \mu v\,|v|^\alpha. \qquad (10)$$

Уравнение (10), в отличие от всех рассмотренных ранее моделей, нелинейно, и его решение выписать, вообще говоря, нельзя (хотя можно провести достаточно детальное изучение системы и в нелинейном случае, в частности, установить, пользуясь приемом, применявшимся для уравнения (6), затухающий характер движения системы). Поэтому ограничимся здесь приближенным анализом поведения системы в двух ее предельных положениях — в окрестности точек $v = 0$ и $r = 0$. Оба положения, очевидно, не могут достигаться одновременно, так как это означало бы, что система покоится.

Если $v(t_0) = 0$ (здесь момент t_0 — один из моментов достижения максимальной амплитуды r_0), то вторым членом в правой части уравнения (10) можно пренебречь по сравнению с первым, и оно принимает вид

$$m\frac{d^2r}{dt^2} = -kr_0.$$

Поскольку рассматривается малая окрестность Δt около момента t_0, то пренебрегается также отклонением Δr в сравнении с r_0. Учитывая, что $v(t_0) = 0$, получаем

$$\Delta r = r - r_0 = -\frac{1}{2}\frac{k}{m}r_0\,(t - t_0)^2,$$

т. е. шарик движется с постоянным (в первом приближении) ускорением, так как на него действует лишь сила натяжения пружины, постоянная в окрестности точки $r = r_0$, а сила трения равна нулю.

При $r(t_0) = 0$ (t_0 — один из моментов прохождения системой начала координат, если, конечно, точка $r = 0$ достигается хотя бы один раз) первый член в правой части мал по сравнению со вторым, и

$$m\frac{d^2r}{dt^2} = -\mu v_0\,|v_0|^\alpha.$$

Здесь также пренебрегается отклонением Δv от значения $v_0 = v(t_0)$ ввиду его малости. Так как $r(t_0) = 0$, то из последнего уравнения следует

$$\Delta r = r = \frac{-\mu v_0\,|v_0|^\alpha}{2m}(t - t_0)^2 + v_0\,(t - t_0).$$

Значит, и в этом положении система испытывает постоянное (в первом приближении) ускорение, определяемое лишь силой трения, поскольку натяжение пружины равно нулю. Данный вывод вполне очевиден и справедлив для всех положений системы, хотя ускорение шарика при $v \neq 0$, $r \neq 0$ определяется уже совместным действием обеих сил. Исключение составляет лишь точка, где $kr_0 = -\mu v_0\,|v_0|^\alpha$, когда правая часть уравнения (10) обращается в нуль и первый член в ускорении системы в этот момент $t = t_0$ равен нулю. Разлагая $r(t)$ в ряд Тейлора в окрестности точки $t = t_0$:

$$r(t) = r(t_0) + \frac{dr}{dt}(t_0)\,(t - t_0) +$$

$$+\frac{1}{2}\frac{d^2r}{dt^2}(t_0)(t-t_0)^2+\frac{1}{6}\frac{d^3r}{dt^3}(t_0)(t-t_0)^3+\ldots,$$

где точками обозначены члены более высокого порядка малости, и принимая во внимание, что $\frac{d^2r}{dt^2}(t=t_0)=0$, находим

$$\Delta r=r(t)-r_0=\frac{dr}{dt}(t_0)(t-t_0)+\frac{1}{6}\frac{d^3r}{dt^3}(t_0)(t-t_0)^3+\ldots,$$

т. е. в главном члене ускорение системы в окрестности точки $t=t_0$ не постоянно, а является линейной функцией времени (см. также упр. 5).

Еще один тип нелинейности может обусловливаться изменяющимися механическими свойствами пружины. Закон Гука действителен, вообще говоря, лишь при малых отклонениях (деформациях) пружины от ненагруженного нейтрального положения. При заметных деформациях пружина, в зависимости от материала, из которого она изготовлена, и величины деформации, может вести себя как «мягкая», и тогда сила натяжения будет меньше, чем полагается по закону Гука (в случае «жесткой» пружины — наоборот). Жесткость пружины в такой ситуации становится функцией координаты, т. е. $k=k(r)$, и уравнение движения принимает вид

$$m\frac{d^2r}{dt^2}=-k(r)\,r, \tag{11}$$

где, разумеется, $k(r)>0$. Например, если $k(r)=k_0/(1+|r|)$, то пружина мягкая. Уравнение (11) нелинейное так же, как и (10), но достаточно очевидны по крайней мере два различия между (10) и (11). В отличие от (10), можно выписать (неявное) решение для (11) с двукратным использованием квадратуры. Кроме того, переписывая (11) в виде

$$m\frac{dr}{dt}\frac{d^2r}{dt^2}=-\frac{dr}{dt}k(r)\,r=\frac{dr}{dt}\frac{d}{dr}\int\limits_0^r k(r')\,r'\,dr'=-\frac{d}{dt}\left(\int\limits_0^r k(r')\,r'\,dr'\right),$$

учитывая, что левая часть этого выражения равна $mv\,dv/dt=m\,dv^2/dt$, и интегрируя его по t, получаем

$$E_к+E_п=m\left(\frac{dr}{dt}\right)^2+\int\limits_0^r k(r')\,r'\,dr'=\text{const}>0. \tag{12}$$

Это означает *консервативность* движения, описываемого моделью (11), или постоянство полной энергии системы. Существование первого интеграла (12) позволяет установить общий со случаем линейной системы факт — колебательный характер изучаемого движения. Действительно, из (12) следует ограниченность функций $v(t)=dr/dt$ и $r(t)$ при любых $t>0$. Решение не имеет предела при $t\to\infty$, так как при $v(t\to\infty)\to v_\infty\neq 0$ это противоречило бы ограниченности функции $r(t)$ при $t\to\infty$, а для $v(t\to\infty)\to v_\infty=0$ невозможно, чтобы

$r(t\to\infty)\to r_\infty\neq 0$ — тогда из (11) следовала бы неограниченность величины $v(t)$ при $t\to\infty$ (случай же $v_\infty=r_\infty=0$ противоречит (12)). Тем самым шарик колеблется. Он неограниченное число раз проходит точку $r=0$ (в противном случае величина $r(t)$ была бы знакоопределенной вместе с ускорением d^2r/dt^2 (см. (11)) и $v\to\infty$, $t\to\infty$).

5. Заключение. Приведенные в этом параграфе построения демонстрируют иерархическую цепочку моделей системы «шарик—пружина», получающихся одна из другой при последовательном отказе от предположений, идеализирующих изучаемый объект. В одних случаях усложнение не вносит ничего нового в поведение системы (постоянная внешняя сила, шарик на двух пружинах), в других ее свойства меняются существенным образом. Путь «от простого к сложному» дает возможность поэтапно изучать все более реалистичные модели и сравнивать их свойства.

Существует и другой путь построения и изучения моделей — «от общего к частному». Из результатов данного параграфа очевидно: достаточно общее уравнение движения системы «шарик—пружина» записывается в виде

$$m\frac{d^2r}{dt^2}=-k(r,t)\,r+F\left(r,t,\frac{dr}{dt}\right),\qquad k>0,$$

где k и F могут быть разнообразными функциями своих аргументов. Опираясь на эту общую модель, можно, проводя соответствующие конкретизации, последовательно получать и изучать более простые модели. Например, зависимость k от r и t отвечает следующему после (5) уравнению и уравнению (11), зависимость F от r, t — наличию внешней силы или силы инерции (уравнения (2), (3), (5)), а от dr/dt — сопротивлению среды (уравнения (6), (8), (10)). Данный подход также широко применяется, в том числе и потому, что позволяет сразу установить некоторые общие свойства объекта, конкретизируя и дополняя их в более частных ситуациях.

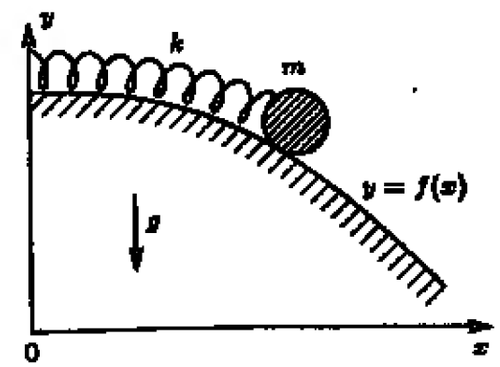


Рис. 14

УПРАЖНЕНИЯ

1. Получите уравнение движения шарика на пружине, перемещающегося по идеальной поверхности с непостоянным наклоном под действием силы натяжения пружины и силы тяжести. Уравнение поверхности: $y=f(x)$, $y'\leqslant 0$ (рис. 14).

2. Подберите величины k_1, k_2, m, r_0, v_0 так, чтобы в системе на рис. 13 отсутствовало соприкосновение шарика с точками крепления.

3. Пользуясь приемом, применявшимся при анализе уравнения (6), покажите, что в случае уравнения (8) движение происходит с затуханием.

4. Выпишите решение уравнения (9) при $k_1<0$ через гиперболические функции и убедитесь в том, что решение уравнения (8) стремится к нулю при $t\to\infty$.

5. Пользуясь разложением в ряд Тейлора функции $r(t)$ в окрестности точки $t=t_0$, где правая часть уравнения (10) равна нулю, найдите величину $d^3r/dt^3(t_0)$.

§ 5. Универсальность математических моделей

Рассмотрим процессы колебаний в объектах различной природы. Покажем, что несмотря на разную сущность объектов им соответствуют одни и те же математические модели.

1. Жидкость в U-образном сосуде. Жидкость занимает часть сосуда U-образной формы, представляющего собой изогнутую трубку радиуса r_0 (рис 15) Масса жидкости M_0, ее плотность ρ_0. Стенки сосуда идеально гладкие, поверхностным натяжением пренебрегается, атмосферное давление и ускорение свободного падения g постоянны

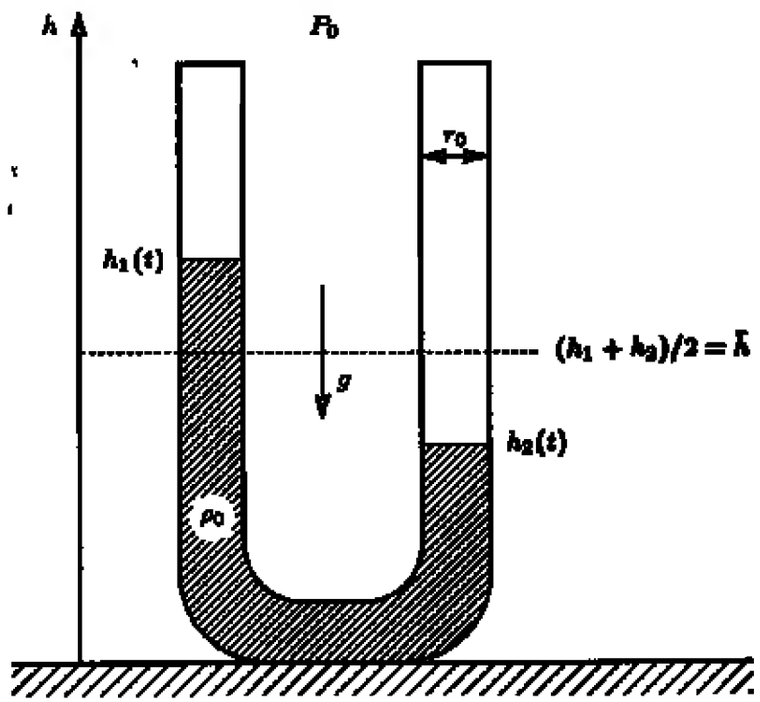


Рис 15

В состоянии равновесия жидкость, очевидно, покоится, ее высота в обоих коленах сосуда одинакова Если ее вывести из равновесия, то начнется движение, характер которого установим с помощью закона сохранения энергии, поскольку в силу сделанных предположений ее потери в системе отсутствуют

Потенциальную энергию системы вычислим через работу, которую необходимо совершить, чтобы переместить ее из состояния равновесия (где $h_1 = h_2$) в положение, изображенное на рис. 15. Она равна

$$E_п = -\int\limits_{\bar{h}}^{h_2} P\,dh_2 = -\int\limits_{\bar{h}}^{h_2} \rho_0 s_0 (h_1 - h)\, g\, dh, \qquad \bar{h} = \frac{h_1 + h_2}{2},$$

где P — вес той части жидкости в левом колене, уровень которой превышает величину h_2. Работа сил атмосферного давления равна нулю,

так как для разных колен соответствующие перемещения направлены в разные стороны.

Неизвестные величины $h_1(t)$ и $h_2(t)$ связаны очевидным соотношением $h_1(t) + h_2(t) = \text{const} > 0$, выражающим постоянство полной длины столба жидкости в сосуде с постоянным сечением. Подставляя последнее равенство в выражение для $E_п$, получаем после интегрирования

$$E_п = -\rho_0 s_0 g\,(-h_2^2(t) + Ch_2(t) + C_1).$$

При вычислении кинетической энергия учтем постоянство сечения трубки и несжимаемость жидкости Это означает, что столб жидкости движется как целое, и ее скорость $v(t)$ одинакова во всех сечениях. Примем за $v(t)$ величину $dh_2(t)/dt$, и тогда

$$E_к = \frac{1}{2} M_0 \left(\frac{dh_2}{dt}\right)^2,$$

а из закона сохранения энергии следует

$$E(t) = E_к(t) + E_п(t) = \frac{M_0}{2}\left(\frac{dh_2}{dt}\right)^2 - \rho_0 s_0 g\,(-h_2^2 + Ch_2 + C_1).$$

Так как $dE/dt = 0$, то, продифференцировав это выражение, получаем

$$M_0 \frac{d^2 h_2}{dt^2} = \rho_0 s_0 g\,(-2h_2 + C),$$

что, с учетом такого же соотношения для величины $h_1(t)$, дает уравнение

$$M_0 \frac{d^2 h}{dt^2} = -\rho_0 s_0 g h = -\pi \rho_0 r_0^2 g h,$$

где $h = (h_2 - h_1)/2$ — отклонение уровня жидкости от положения равновесия Оно, с точностью до обозначений, полностью совпадает с уравнением (1) § 4 для системы «шарик—пружина» (в данном случае аналогом шарика служит столб жидкости, а роль пружины играет тяготение)

Последовательный отказ от идеализации объекта дает более полные его модели (как и в § 4) Например, учет силы поверхностного натяжения, равной $\sigma_0 2\pi r_0$ (σ_0 — коэффициент поверхностного натяжения) и всегда направленной против движения жидкости, приводит к уравнению вида (7) § 4 для величины h (см также упр 1)

2. Колебательный электрический контур. Это устройство представляет собой конденсатор, соединенный проводами с индуктивной катушкой В момент $t = 0$ цепь замыкается, и заряд с обкладок конденсатора начинает распространяться по цепи (рис 16)

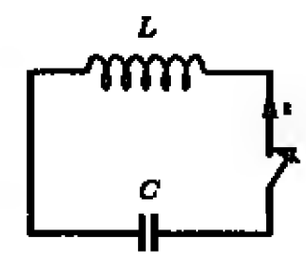


Рис. 16

Сопротивление проводов будем считать равным нулю, емкость конденсатора равна C, индуктивность катушки

L. Для изменяющейся со временем величины $q(t)$, где $q(t)$ — заряд на обкладках конденсатора, необходимо получить соответствующее уравнение. Очевидно, что ток $i(t)$ и напряжение $v(t)$ также являются функциями времени.

По физическому смыслу величины C в любой момент времени имеем равенство $v(t) = q(t)\,C$ (емкость равна величине заряда, который необходимо поместить на обкладки конденсатора для увеличения разности потенциалов между ними на единицу).

Так как электрическое сопротивление в цепи отсутствует, то падения напряжения на проводах нет, и разность потенциалов $v(t)$, существующая на конденсаторе, подается непосредственно на катушку. При переменном токе в катушке возникает электродвижущая сила самоиндукции, равная $\varepsilon = -L\,di/dt$. Закон Ома для цепи в отсутствие сопротивления выглядит следующим образом:

$$v(t) = -\varepsilon(t),$$

или

$$q(t)\,C = -\varepsilon(t) = L\,di/dt.$$

Так как по определению $i = -dq/dt$ (при убывании заряда на конденсаторе ток в цепи увеличивается, и наоборот), то из последнего соотношения получаем уравнение

$$L\,\frac{d^2q}{dt^2} = -Cq,$$

описывающее процесс колебаний величины $q(t)$ (а следовательно, и величин $i(t)$, $v(t)$) в простейшем электрическом контуре, тождественное (1) § 4. В системе «емкость—индуктивность» колебания происходят так же, как и в системе «шарик—пружина» (и так же усложняются соответствующие модели при учете дополнительных процессов — см. упр. 2).

3. Малые колебания при взаимодействии двух биологических популяций. Пусть на одной и той же территории проживают две биологические популяции с численностями $N(t)$ и $M(t)$, причем первая растительноядная, а вторая употребляет в пищу представителей первой популяции.

Скорость изменения $N(t)$ складывается из определяемой по первому члену в правой части формулы (10) § 1 скорости прироста благодаря рождаемости (эффект насыщения не учитывается; ср. с (12) § 1) и из скорости убывания благодаря соседству со второй популяцией:

$$\frac{dN}{dt} = (\alpha_1 - \beta_1 M)\,N, \tag{1}$$

где $\alpha_1 > 0, \beta_1 > 0$, член $\beta_1 M N$ описывает вынужденное убывание (естественной смертностью популяции пренебрегаем).

Численность второй популяции растет тем быстрее, чем больше численность первой популяции, а при ее отсутствии уменьшается со скоростью, пропорциональной численности $M(t)$ (тем самым ее рождаемость не учитывается, как и эффект насыщения):

$$\frac{dM}{dt} = (-\alpha_2 + \beta_2 N)\,M, \tag{2}$$

где $\alpha_2 > 0$, $\beta_2 > 0$

Очевидно, что система находится в равновесии при $M_0 = \alpha_1/\beta_1$ и $N_0 = \alpha_2/\beta_2$, когда $dN/dt = dM/dt = 0$. Рассмотрим малые отклонения системы от равновесных значений, т. е. представим решение в виде $N = N_0 + n$, $M = M_0 + m$, $n \ll N_0$, $m \ll M_0$. Подставляя N и M в уравнения (1), (2), получим, отбрасывая члены более высокого порядка малости,

$$\frac{dn}{dt} = -\beta_1 N_0 m, \tag{3}$$

$$\frac{dm}{dt} = -\beta_2 M_0 n. \tag{4}$$

Дифференцируя (3) по t и подставляя в полученное уравнение функцию dm/dt, определяемую из (4), придем к уравнению

$$\frac{d^2n}{dt^2} = -\alpha_1\alpha_2 n,$$

аналогичному по форме уравнению (1) § 4. Следовательно, в системе происходят малые колебания численности с частотой $\omega = \sqrt{\alpha_1\alpha_2}$, зависящей только от коэффициентов рождаемости и смертности α_1 и α_2.

Заметим, что величина $m(t)$ подчиняется такому же уравнению, причем если отклонение $n(t)$ равно нулю в начальный момент $t = 0$, то $m(t = 0)$ имеет максимальную амплитуду, и наоборот (см. решение уравнения колебаний (6) § 2) Эта ситуация, когда численности $n(t)$ и $m(t)$ находятся в противофазе, воспроизводится для всех моментов $t_i = iT/4$, $i = 1, 2, \ldots$ (T — период колебаний) и отражает запаздывание реакции численности одной популяции на изменение численности другой (см. также упр 3)

4. Простейшая модель изменения зарплаты и занятости. Рынок труда, на котором взаимодействуют работодатели и наемные рабочие, характеризуется зарплатой $p(t)$ и числом занятых $N(t)$. Пусть на нем существует равновесие, т. е. ситуация, когда за плату $p_0 > 0$ согласны работать $N_0 > 0$ человек. Если по каким-то причинам это равновесие нарушается (например, по возрасту часть работников уходит на пенсию либо у предпринимателей возникают финансовые трудности), то функции $p(t)$ и $N(t)$ отклоняются от значений p_0, N_0.

Будем считать, что работодатели изменяют зарплату пропорционально отклонению численности занятых от равновесного значения. Тогда

$$\frac{dp}{dt} = -\alpha_1\,(N - N_0), \qquad \alpha_1 > 0.$$

Предположим, что число работников увеличивается или уменьшается также пропорционально росту или уменьшению зарплаты относительно значения p_0, т. е.

$$\frac{dN}{dt} = \alpha_2 (p - p_0), \qquad \alpha_2 > 0.$$

Дифференцируя первое уравнение по t и исключая из него с помощью второго уравнения величину N, приходим к стандартной модели колебаний

$$\frac{d^2(p - p_0)}{dt^2} = -\alpha_1 \alpha_2 (p - p_0)$$

заработной платы относительно положения равновесия (аналогично и для величины $N(t)$). Из первого интеграла этого уравнения

$$\alpha_1 (N - N_0)^2 + \alpha_2 (p - p_0)^2 = \text{const} > 0$$

видно, что в некоторые моменты $t = t_i$, $i = 1, 2, \ldots$, когда $p = p_0$ (т. е. зарплата становится равной равновесному значению), имеем $N > N_0$, т. е. число занятых больше равновесного, а при $N = N_0$ получаем $p > p_0$, т. е. зарплата превышает равновесную. В эти моменты фонд заработной платы, равный pN, превышает равновесное значение $p_0 N_0$ (или меньше его), если при подходе к моменту t_i выполнено $p > p_0$ или $N > N_0$ (и наоборот). Но в среднем за период колебаний величина pN равна $p_0 N_0$ (упр. 4).

5. Заключение. Построенные в данном параграфе модели в одних случаях основаны на точно известных законах (пп. 1, 2), в других — на наблюдаемых фактах и либо на аналогиях (п. 3), либо на правдоподобных представлениях о характере объекта (п. 4). Хотя и сущность рассматривавшихся явлений, и подходы к получению отвечающих им моделей совершенно различны, построенные модели оказались идентичны друг другу. Это свидетельствует о важнейшем свойстве математических моделей — их *универсальности*, — широко используемом при изучении объектов самой разнообразной природы.

УПРАЖНЕНИЯ

1. Пусть в задаче об U-образном сосуде левое колено имеет переменное сечение, т. е. $r = r_0(h)$. Покажите, применяя второй закон Ньютона и предполагая отсутствие горизонтальной компоненты у скорости жидкости, что для величины h получается уравнение вида (11) § 4.

2. Вводя в LC-контур сопротивление R и используя закон Ома, убедитесь в том, что модель колебаний в LCR-контуре аналогична уравнению (8) § 4.

3. Сведите нелинейную систему (1), (2) к уравнению второго порядка и покажите, что она (как и ее линейный аналог (3), (4)) имеет первый интеграл.

4. Пользуясь формулой (6) § 2 для общего решения уравнения колебаний, покажите, что среднее значение фонда заработной платы pN (п. 4) за период колебаний равно равновесному.

§ 6. Некоторые модели простейших нелинейных объектов

Обсудим происхождение нелинейности и рассмотрим некоторые ее последствия, проявляющиеся в поведении изучаемых объектов. Проиллюстрируем неизбежность применения численных методов для их анализа.

1. О происхождении нелинейности. Как уже отмечалось в п. 5 § 1, линейные модели подчиняются принципу суперпозиции. В этом случае, находя частные решения и суммируя их, как правило, удается построить и общее решение (типичные примеры — формула (6) § 2 и формула для общего решения уравнения (3) § 4 в моделях колебаний).

Для нелинейных моделей принцип суперпозиции неприменим, и общее решение можно найти лишь в редких случаях. Отдельные же частные решения нелинейных уравнений могут не отражать характер поведения объекта в более общей ситуации.

Источниками нелинейности могут быть многие причины. Фундаментальные законы природы — закон тяготения и закон Кулона — изначально нелинейны (квадратичная зависимость силы взаимодействия между массами или зарядами), и потому основанные на них модели, вообще говоря, также нелинейны. Свой вклад в нелинейность моделей вносят более сложная геометрия явления (см. упр. 1 § 4 и упр. 1 § 5), различные внешние воздействия (см. уравнение (10) § 4) и, конечно же, изменение характера взаимодействия в самом объекте при изменении его состояния (эффект насыщения в моделях популяций, меняющаяся жесткость пружины).

В сущности, реальным явлениям отвечают только нелинейные модели, а линейные справедливы лишь при описании незначительных изменений величин, характеризующих объект.

2. Три режима в нелинейной модели популяции. В отличие от модели Мальтуса (10) § 1 и модели (12) § 1 коэффициент рождаемости будем считать зависящим от численности популяции $N(t)$, т. е. $\alpha = \alpha(N)$. Коэффициент смертности β также зависит от N. Уравнение динамики популяции

$$\frac{dN}{dt} = (\alpha(N) - \beta(N))\, N \tag{1}$$

нелинейно благодаря изменению характеристик взаимодействия внутри популяции при изменении ее состояния.

Положим для определенности $\beta(N) = \beta_0 = \text{const}$, $\alpha(N) = \alpha_0 N$, т. е. рождаемость пропорциональна численности (например, потому что члены популяции заинтересованы в ее росте). Тогда уравнение (1) преобразуется к виду

$$\frac{dN}{dt} = \alpha_0 N^2 - \beta_0 N \tag{2}$$

с квадратичной нелинейностью (характерной также для некоторых химических реакций). Рассмотрим поведение функции $N(t)$ при различных начальных численностях $N(0) = N_0$ (рис. 17).

а) При $N_0 < N_{кр} = \beta_0/\alpha_0$ численность монотонно уменьшается со временем, стремясь к нулю при $t \to \infty$. Решение дается формулой, аналогичной формуле для решения уравнения (12) § 1, где t заменяется на $-t$ (обратная логистическая кривая; ср. с рис. 7, § 1).

б) При критическом значении $N_0 = N_{кр}$ численность популяции не зависит от времени.

в) При $N_0 > N_{кр}$ характер решения принципиально изменяется по сравнению со случаями а) и б): численность растет со временем, причем настолько быстро, что обращается в бесконечность за конечное время $t = t_f$. Величина t_f тем меньше, чем больше N_0 (см. упр. 1).

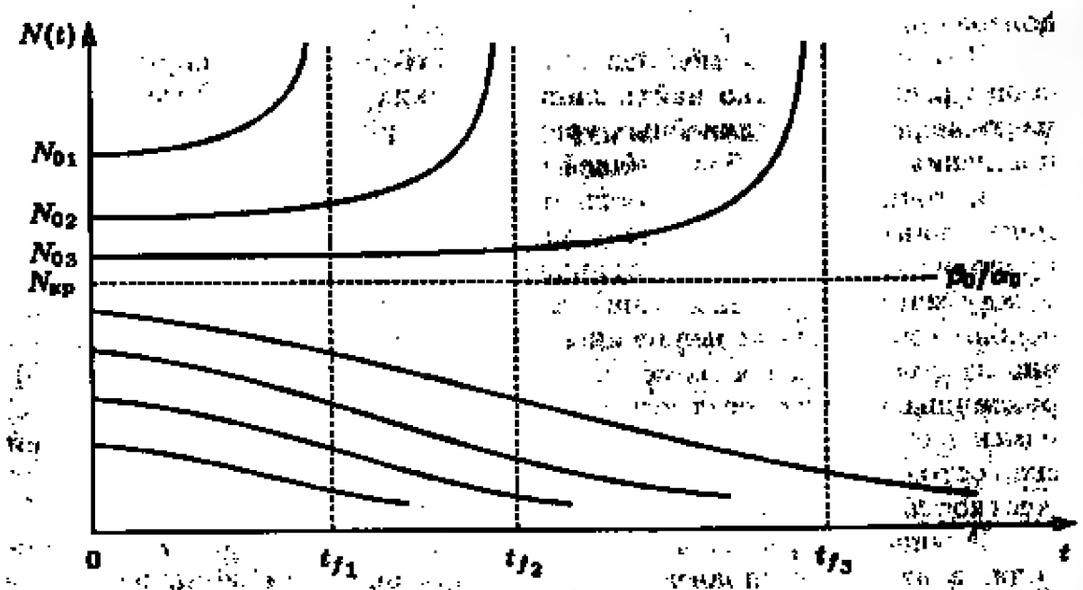


Рис. 17

Нелинейность уравнения (2) порождает большое разнообразие эффектов, содержащихся даже в простейшей модели: три возможных режима изменения численности со временем; неустойчивость режима б) — при малых отклонениях в область а) или в) решение удаляется от линии $N_{кр} = \beta_0/\alpha_0$; сильную чувствительность функции $N(t)$ к начальным данным N_0; наконец, катастрофический рост численности популяции за конечное время при $N_0 > N_{кр}$.

Заметим, что последнее свойство не частный результат, а имеет место для любых моделей вида

$$\frac{dN}{dt} = F(N), \qquad t > 0, \quad N(0) > 0, \quad F(N) > 0,$$

если при больших N функция $F(N)$ растет быстрее первой степени N, точнее, если для $F(N)$ справедлив критерий

$$\int\limits_{N(0)}^{\infty} \frac{dN}{F(N)} < \infty,$$

получающийся непосредственным интегрированием уравнения.

3. Влияние сильной нелинейности на процесс колебаний. Уравнение колебаний

$$m\frac{d^2r}{dt^2} = -k(r)\,r, \tag{3}$$

где функция $k(r) > 0$ описывает жесткость пружины, — одно из относительно немногих нелинейных уравнений, для которого можно выписать общее решение. Вводя величину скорости $v = dr/dt$, перепишем (3) в виде

$$m\frac{dv}{dt} = -k(r)\,r, \qquad \frac{dr}{dt} = v;$$

деля первое из этих уравнений на второе, получим нелинейное уравнение первого порядка

$$m\frac{dv}{dr} = -\frac{k(r)\,r}{v}. \tag{4}$$

Разделяя в (4) переменные:

$$mv\,dv = -k(r)\,r\,dr,$$

и дважды интегрируя последнее уравнение, находим

$$v^2 = \left(\frac{dr}{dt}\right)^2 = -2\int\limits_0^r k'(r')\,r'\,dr' + C, \qquad k'(r) = \frac{k(r)}{m},$$

$$\frac{dr}{dt} = \pm\sqrt{C - 2\int\limits_0^r k'(r')\,r'\,dr'},$$

$$t = \pm\int\limits_0^r d\bar{r}\left(\sqrt{C - 2\int\limits_0^{\bar{r}} k'(r')\,r'\,dr'}\right)^{-1} + C_1, \tag{5}$$

где в неявно выписанном общем решении (5) константы C, C_1 можно определить, зная начальные данные.

Заметим, что данная процедура по отношению к уравнениям вида (10) § 4 неприменима (переменные не разделяются), и найти их общее решение этим способом уже нельзя.

В линейном случае ($k(r) = k_0$) интегральные кривые уравнения (4) представляют собой концентрические круги с центром в начале координат, радиус которых определяется начальной энергией системы и «движение» по которым описывает периодический во времени процесс колебаний (рис. 18).

Рассмотрим теперь сильно нелинейную систему, в которой пружина ведет себя как «сверхмягкая», например, $k(r) = 1/(r^2 + \alpha)$, $\alpha > 0$. В предельном случае $\alpha = 0$ уравнение (4) принимает вид

$$m\frac{dv}{dr} = -\frac{1}{vr},$$

и его решение принципиально отличается от решения (4) (см. рис. 19) тем, что энергия не сохраняется и, более того, неограниченно растет

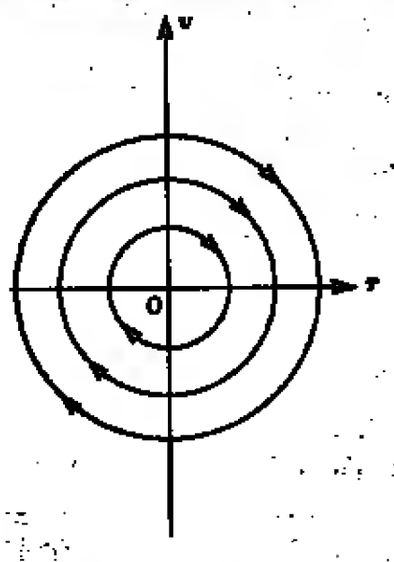


Рис. 18

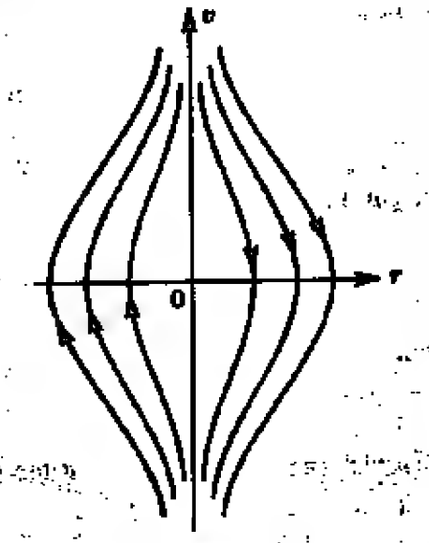


Рис. 19

при $r \to \pm 0$. При ослаблении нелинейности процесс колебаний приобретает обычный характер (упр. 2).

4. О численных методах. Рассмотренные здесь примеры достаточно убедительно свидетельствуют о неизбежности применения численных методов для моделирования нелинейных объектов из-за явной недостаточности чисто теоретических подходов и сложного, разнообразного поведения характеризующих эти объекты величин. Впрочем, этот вывод справедлив и для линейных моделей, содержащих большое число неизвестных величин, независимых переменных, параметров и имеющих сложную пространственную структуру. Для построения соответствующих численных моделей широко используются методы, подходы и приемы, разрабатываемые при создании исходных моделей, и возникают свои специфические проблемы, требующие глубокого изучения.

Поясним последнее утверждение простым примером. Для уравнения (10) § 1

$$\frac{dN}{dt} = (\alpha - \beta)N = \gamma N, \qquad t > 0, \quad N(0) = N_0,$$

где для определенности $(\alpha - \beta) > 0$, вполне логично предложить следующую численную схему (разбив ось t на равные отрезки величины $\tau = t_{i+1} - t_i$, $i = 0, 1, 2, \ldots$; $t_0 = 0$ и заменив производную на конечную

разность):

$$\frac{N_{i+1} - N_i}{\tau} = \gamma N_i, \qquad i = 0, 1, \ldots; \quad N(t_0) = N_0. \tag{6}$$

Из (6) получаем

$$N_{i+1} = (\tau\gamma + 1)N_i,$$

что дает для его решения

$$N_1 = (1 + \tau\gamma)N_0, \qquad N_2 = (1 + \tau\gamma)^2 N_0,$$

$$N_i = (1 + \tau\gamma)^i N_0 = (1 + \tau\gamma)^{t/\tau} N_0,$$

т. е. при $t \to \infty$ решение (6) может отличаться от искомого сколь угодно сильно. Следовательно, для получения нужной точности необходимо должным образом выбирать шаг τ в зависимости от величины отрезка интегрирования T (упр. 3).

УПРАЖНЕНИЯ

1. Пользуясь заменой, примененной при анализе уравнения (8) § 4, найдите решение уравнения (2) при $N_0 > N_{кр}$ и вычислите величину t_f через N_0, α_0, β_0.

2. Найдите ограничение на рост функции $k(r) \to \infty$, $r \to 0$ в уравнении (3), при выполнении которого система «шарик—пружина» была бы консервативной, т. е. сохранялась бы ее полная энергия.

3. Используя представление числа e в виде соответствующего предела, покажите, что для заданных величин γ, N_0, T решение уравнения (6) стремится при $\tau \to 0$ к решению исходной задачи.

В заключение, основываясь на материале данной главы, выделим ряд тем, являющихся ключевыми для развития и применения методологии математического моделирования. К ним относятся: вопросы идеализации исходного объекта и формулировка соответствующих предположений; применение как строгих процедур (фундаментальные законы, вариационные принципы), так и метода аналогий и других подходов к построению математических моделей (в том числе и трудноформализуемых); методы качественного исследования нелинейных моделей; построение эффективных вычислительных алгоритмов, реализующих модели. Эти вопросы, наряду с описанием некоторых актуальных приложений, составляют основное содержание последующих глав.

Библиография к главе I: [7, 16, 25, 40, 47, 57, 60, 66, 73, 76–79, 81, 83, 84].

ГЛАВА II

ПОЛУЧЕНИЕ МОДЕЛЕЙ ИЗ ФУНДАМЕНТАЛЬНЫХ ЗАКОНОВ ПРИРОДЫ

§ 1. Сохранение массы вещества

На основе составления баланса массы вещества и некоторых дополнительных соображений построим модели потока невзаимодействующих частиц и движения грунтовых вод в пористой среде. Опишем ряд свойств полученных моделей и обсудим их возможные обобщения.

1. Поток частиц в трубе. В цилиндрической трубе с поперечным сечением S (рис. 20) движутся частицы вещества (пылинки, электроны). Скорость их движения $u(t) > 0$ вдоль оси x, вообще говоря, изменяется со временем. Например, заряженные частицы могут ускоряться или замедляться под действием электрического поля. Для построения простейшей модели рассматриваемого движения введем следующие предположения:

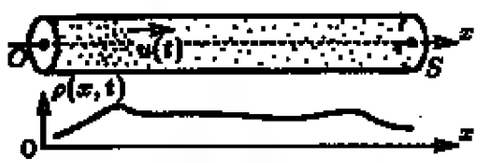


Рис. 20

а) частицы между собой не взаимодействуют (не сталкиваются, не притягиваются и т. д.). Для этого, очевидно, плотность частиц должна быть достаточно малой (в этом случае заряженные частицы не только не сталкиваются, но и не оказывают друг на друга влияния из-за большого расстояния между ними);

б) начальная скорость всех частиц, находящихся в одном и том же поперечном сечении с координатой x, одинакова и направлена вдоль оси x;

в) начальная плотность частиц также зависит только от координаты x;

г) внешние силы, действующие на частицы, направлены вдоль оси x.

Предположение а) означает, что скорость частиц может изменяться лишь под действием внешних сил, предположения б)–г) обеспечивают одномерность процесса переноса, т. е. зависимость искомой плотности потока частиц только от координаты x и времени $t \geqslant 0$.

Итак, по заданной начальной плотности $\rho(x, t = 0) = \rho_0(x)$ необходимо найти плотность частиц $\rho(x,t)$ в любой момент времени для

любых x (скорость движения $u(t)$ задана). Прибегнем к закону сохранения массы, подсчитав баланс вещества в малом элементе трубы от x до $x + dx$ за время dt (рис. 21). Слева в элементарный объем входит количество вещества с массой, равной

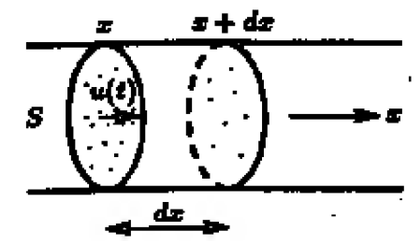


Рис. 21

$$Su(t)\,dt\,\rho(x, t + \xi\,dt), \qquad 0 \leqslant \xi \leqslant 1,$$

где $Su(t)\,dt$ — объем вошедшего за промежуток времени dt вещества. Через правое сечение элемента за то же время выходит масса, равная

$$-Su(t)\,dt\,\rho(x + dx, t + \bar{\xi}\,dt), \qquad \bar{\xi} \neq \xi, \quad 0 \leqslant \bar{\xi} \leqslant 1,$$

т. е. суммарное изменение массы равно

$$dm = Su(t)\,(\rho(x, t + \xi\,dt) - \rho(x + dx, t + \bar{\xi}\,dt))\,dt.$$

В силу малости промежутка dt скорость $u(t)$ считается постоянной. Величины $\rho(x, t + \xi\,dt)$ и $\rho(x + dx, t + \bar{\xi}\,dt)$ — средние по времени значения плотности в сечениях x и $x + dx$.

Другой способ подсчета изменений в фиксированном объеме $S\,dx$ очевиден из смысла величины $\rho(x,t)$:

$$dm = S\,dx\,(\rho(x + \eta\,dx, t + dt) - \rho(x + \bar{\eta}\,dx, t)), \qquad 0 < \eta, \bar{\eta} < 1,$$

где $\rho(x + \eta\,dx, t + dt)$ и $\rho(x + \bar{\eta}\,dx, t)$ — средние по пространству значения плотности в моменты t и $t + dt$.

Приравнивая оба полученные для dm выражения и устремляя dx и dt к нулю, приходим к уравнению для $\rho(x,t)$, отвечающему закону сохранения массы,

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho}{\partial x}\,u(t) = 0, \qquad -\infty < x < \infty, \quad t > 0, \tag{1}$$

с начальным условием

$$\rho(x, 0) = \rho_0(x), \qquad -\infty < x < \infty. \tag{2}$$

Величина ρu (*поток вещества*, или *поток массы*) равна количеству вещества, проходящему в единицу времени через единичную поверхность поперечного сечения трубы. Как видно из (1), скорость изменения плотности вещества со временем в любом сечении определяется «скоростью» изменения потока вещества по координате x. Схожим свойством обладают многие модели, отвечающие законам сохранения и описывающие совсем другие процессы.

В случае постоянной скорости $u(t) = u_0$ приходим к простейшему линейному уравнению в частных производных

$$\frac{\partial \rho}{\partial t} + u_0\,\frac{\partial \rho}{\partial x} = 0, \qquad -\infty < x < \infty, \quad t > 0. \tag{3}$$

Его общее решение нетрудно найти, приняв во внимание, что уравнение (3) имеет характеристики — линии $x = u_0 t + C$, на которых значения искомой функции постоянны во времени, т. е. $\rho(x = u_0 t + C, t) = \rho_c$, или, в эквивалентной записи,

$$\rho(x,t) = \rho(x + u_0(t - t_0), t_0), \qquad t - t_0 \geqslant 0.$$

Выбирая $t_0 = 0$, получим

$$\rho(x,t) = \rho(\xi) = \rho(x + u_0 t). \tag{4}$$

Интеграл (4) и является общим решением уравнения (3). Из формулы (4) и начальных данных (2) легко найти искомую функцию, причем она зависит не по отдельности от переменных x, t, а от их комбинации $\xi = x + u_0 t$ (*бегущая волна*). Пространственный профиль плотности без искажений переносится вдоль потока (рис. 22) с постоянной скоростью (уравнение (3) называют также *уравнением переноса*). Это основное свойство решения уравнения (3) несколько модифицируется в случае, когда скорость частиц зависит от времени (см. упр. 1) — профиль плотности переносится за равные промежутки времени на разные расстояния. Если же по каким-то причинам скорость потока зависит от плотности ($u = u(\rho)$), то уравнение (1) становится нелинейным, и поведение его решения может иметь качественно иной характер (см. п. 7 § 4).

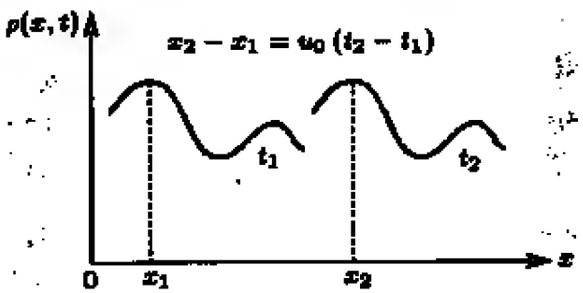


Рис. 22

2. Основные предположения о гравитационном режиме течения грунтовых вод. Пористая среда представляет собой пласт водопроницаемого материала (песок, глина), ограниченного снизу грунтом, не пропускающим воду (гранит), а сверху — поверхностью земли (рис. 23). Если из-за интенсивной работы артезианских скважин или в результате обильных осадков уровень воды в каком-либо месте слоя изменяется, то под действием силы тяжести начинается движение жидкости, выравнивающее ее свободную поверхность.

Для описания этого процесса введем прежде всего ряд предположений:

1) вода рассматривается как несжимаемая жидкость с постоянной плотностью ρ;

2) толщина пласта много меньше его ширины и длины;

3) подстилающая поверхность не имеет разрывов и изломов, задающая ее известная функция $H(x,y)$ — достаточно гладкая функция своих аргументов;

4) свободная поверхность воды $h = h(x,y,t)$ плавно меняется с изменением координат x, y;

5) грунтовые воды нигде не выходят на поверхность земли, причем на свободной поверхности жидкости давление постоянно;

6) грунт однороден, т. е. его физико-механические свойства не зависят от аргументов x, y, z.

Первое предположение вполне естественно, поскольку в рассматриваемом процессе не могут достигаться давления, способные заметно изменять плотность воды. Остальные предположения упрощающие. Например, второе предположение (тонкий пласт) означает, что течение жидкости двумерное и все его характеристики не зависят от координаты z, последние два предположения позволяют построить модель, единообразную во всех точках грунта, и т. д. Вместе с тем предположения 1)–6) отнюдь не выхолащивают суть процесса, так как они выполняются в большом количестве реальных ситуаций.

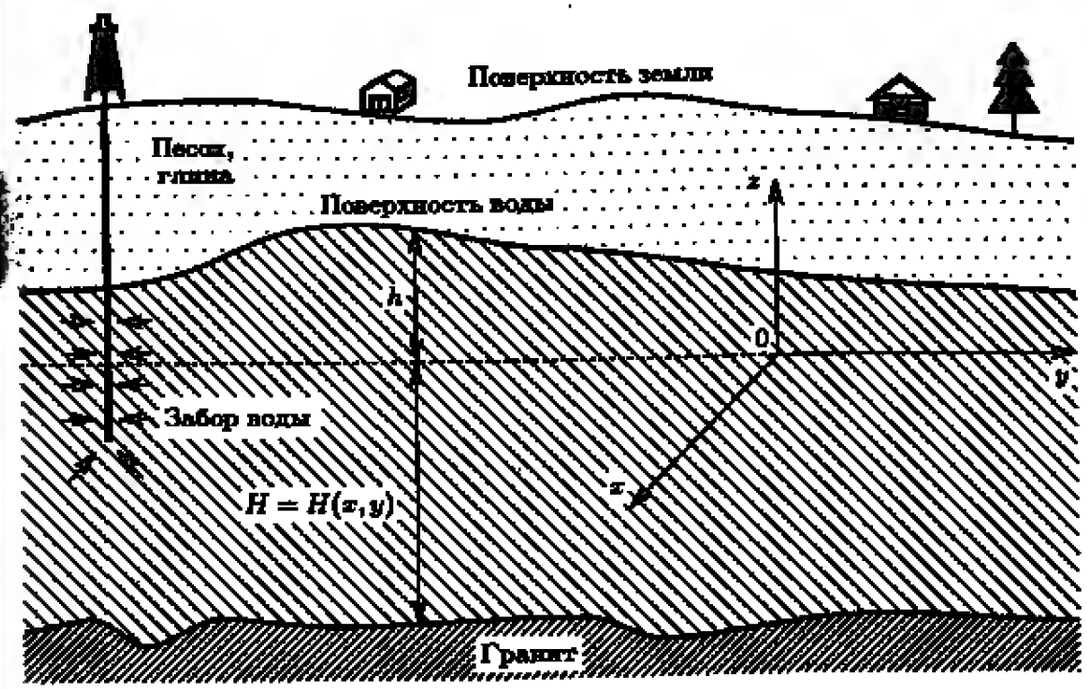


Рис. 23

3. Баланс массы в элементе грунта. Выделим в пласте элементарный объем, образующийся в результате пересечения вертикальной призмы $ABCD$ подстилающей и свободной поверхностями грунта. Поскольку размеры призмы dx и dy малы, а функции H и h гладкие (предположения 3), 4)), то получившееся тело с хорошей степенью точности можно считать параллелепипедом. Введем неизвестные функции $v = v(x,y,t)$ и $u = u(x,y,t)$ — составляющие скорости жидкости вдоль осей x, y (рис. 24).

Подсчитаем количества жидкости, входящей в параллелепипед и выходящей из него за промежуток времени dt.

Через грань DC в элемент грунта входит масса воды, равная объему прошедшей через нее жидкости, умноженному на плотность ρ, т. е. величина

$$\rho u (H + h)\, dy\, dt,$$

а через грань AB выходит масса воды

$$\rho u (H + h)\, dy\, dt + \left\{ \frac{\partial}{\partial x} [\rho u (H + h)]\, dx \right\} dy\, dt.$$

В этом выражении по сравнению с предыдущим добавляется член, описывающий приращение функции $\rho u (H + h)$ при переходе от плоскости

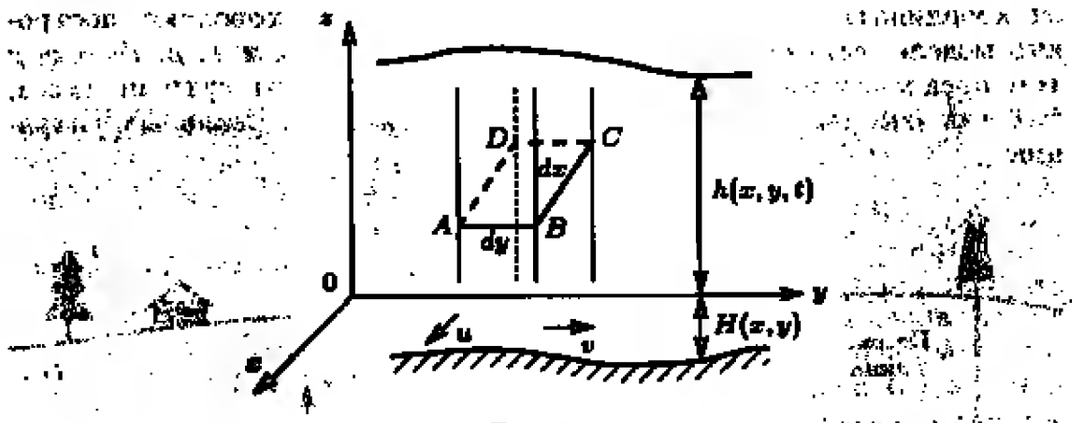


Рис. 24

x к плоскости $x + dx$. Сама же величина $\rho u (H + h)$, как и в п. 1, имеет смысл потока массы (вещества).

Итак, при движении жидкости вдоль оси x в элементе dx грунта накапливается масса

$$-\frac{\partial}{\partial x} [\rho u (H + h)]\, dx\, dy\, dt.$$

Проводя аналогичные рассуждения для граней AD и BC, получаем изменение массы воды за счет ее движения вдоль оси y:

$$-\frac{\partial}{\partial y} [\rho v (H + h)]\, dx\, dy\, dt.$$

Поскольку вдоль оси z в элемент грунта жидкость не втекает и не вытекает из него (снизу — подстилающий пласт, а через свободную поверхность нет потока вещества), то суммарное изменение массы воды в элементе грунта равно

$$-\left\{ \frac{\partial}{\partial x} [\rho u (H + h)] + \frac{\partial}{\partial y} [\rho v (H + h)] \right\} dx\, dy\, dt. \qquad (5)$$

Общее количество жидкости в параллелепипеде равно его объему, умноженному на плотность ρ и на коэффициент пористости $m < 1$ (так как часть объема занята грунтом):

$$m\rho (H + h)\, dx\, dy.$$

Изменение массы воды в элементе за время dt, очевидно, равно

$$\left\{ \frac{\partial}{\partial t} [\rho m (H + h)]\, dx\, dy \right\} dt.$$

Учитывая, что $\partial H/\partial t \equiv 0$, $\partial \rho/\partial t \equiv 0$, из последнего выражения получаем

$$m\rho \frac{\partial h}{\partial t}\, dx\, dy\, dt \qquad (6)$$

и, приравнивая (5) к (6), приходим к *уравнению неразрывности*, выражающему закон сохранения массы в рассматриваемом процессе:

$$m\rho \frac{\partial h}{\partial t} = -\frac{\partial}{\partial x} [\rho u (H + h)] - \frac{\partial}{\partial y} [\rho v (H + h)]. \qquad (7)$$

В уравнении (7) скорость изменения рассматриваемой величины (в данном случае массы) со временем определяется дивергенцией потока этой величины — свойство, характерное для многих моделей, получаемых из законов сохранения (ср. с уравнением (1)).

С учетом того, что $\partial \rho/\partial x \equiv 0$, $\partial \rho/\partial y \equiv 0$, уравнение (7) переписывается в более простой форме:

$$m \frac{\partial h}{\partial t} = -\frac{\partial}{\partial x} [u (H + h)] - \frac{\partial}{\partial y} [v (H + h)]. \qquad (8)$$

4. Замыкание закона сохранения массы. Уравнение (8) содержит три неизвестных величины — h, u, v. Следовательно, для замыкания модели необходимо привлечь какие-то дополнительные соображения о характере процесса. Их дает полуэмпирический *закон Дарси*

$$u = -\mu \frac{\partial p}{\partial x}, \qquad v = -\mu \frac{\partial p}{\partial y}, \qquad (9)$$

где $p(x, y, t)$ — давление в жидкости, $\mu > 0$ — коэффициент, определяемый свойствами грунта. Согласно закону Дарси компоненты скорости течения жидкости пропорциональны соответствующим компонентам градиента давления. Заметим, что по своему физическому смыслу градиент давления — это сила (отнесенная к единице объема). В то же время по второму закону Ньютона действующая на тело сила пропорциональна его ускорению, а не скорости, как в законе Дарси. Однако данное противоречие кажущееся, так как при течении через грунт (фильтрации) жидкость преодолевает сопротивление его частиц, в отличие от свободного течения (ср. с уравнением движения жидкости в § 4).

В формулах (9) используется новая неизвестная величина — давление жидкости. Ее связь с уже введенными величинами нетрудно найти, приняв предположение о медленном и почти горизонтальном течении воды. Тогда динамической составляющей давления можно пренебречь и вычислять его по чисто гидростатическому закону как давление, создаваемое столбом жидкости:

$$p(x, y, z, t) = \rho g (h(x, y, t) - z) + \text{const},$$

где const — давление на поверхности жидкости (например, атмосферное), g — ускорение свободного падения.

Подставляя последнюю формулу в (9), получаем

$$u = -\mu\rho g\,\frac{\partial h}{\partial x}, \qquad v = -\mu\rho g\,\frac{\partial h}{\partial y}, \tag{10}$$

и, используя (10) в уравнении неразрывности (8), окончательно приходим к уравнению движения грунтовых вод

$$\frac{\partial h}{\partial t} = k\,\frac{\partial}{\partial x}\left[(H(x,y)+h)\,\frac{\partial h}{\partial x}\right] + k\,\frac{\partial}{\partial y}\left[(H(x,y)+h)\,\frac{\partial h}{\partial y}\right], \tag{11}$$

$$k = \frac{\mu\rho g}{m},$$

или к *уравнению Буссинеска*, содержащему лишь одну неизвестную функцию $h(x,y,t)$.

5. О некоторых свойствах уравнения Буссинеска. Уравнение (11) нестационарное (искомая функция h зависит от t), двумерное (h зависит от x и y), относящееся к параболическому типу. Оно неоднородное, так как функция H зависит от x, y, и нелинейное, поскольку в его правой части присутствуют члены вида $(hh_x)_x$ и $(hh_y)_y$. В сравнении с уравнением (1) уравнение Буссинеска — гораздо более сложный математический объект. В силу нелинейности его общее решение не может быть найдено аналитически, однако относительно нетрудно получить некоторые вполне содержательные частные решения (см. упр. 2), которые служат также тестами при разработке численных методов для уравнения (11).

Для построения завершенной модели движения грунтовых вод необходимо знать входные данные для уравнения (11): форму подстилающей поверхности $H(x,y)$, коэффициент k и краевые условия, задающие функцию h в начальный момент времени и на границах пласта (и, быть может, в некоторых выделенных областях пласта, например, на артезианской скважине). Более подробно формулировка краевых условий для уравнений параболического типа обсуждается в § 2. Здесь отметим только, что простейшим вариантом формулировки краевых условий для уравнения (11) является задание лишь начального условия — функции $h(x,y,t)$ в момент $t=0$:

$$h(x,y,t=0) = h_0(x,y), \qquad -\infty < x < \infty, \quad -\infty < y < \infty.$$

Такая постановка отвечает *задаче Коши* для уравнения (11), решаемого, естественно, также в области $-\infty < x < \infty$, $-\infty < y < \infty$. В задаче Коши по известному распределению уровня грунтовых вод h_0 находится функция h для всех $t > 0$.

Рассмотрение пласта бесконечных размеров, конечно же, идеализация. Однако если изучается течение в небольшой центральной области пласта на относительно небольшом промежутке времени, то влиянием границ пласта можно пренебречь, и решение задачи Коши описывает вполне реальный процесс.

Отметим также, что некоторые краевые условия были фактически неявно введены в модель при выводе уравнения Буссинеска. Предположение о непроницаемости пласта было использовано при получении

уравнения баланса, а без предположения 5) о «зазоре» между поверхностью земли и поверхностью грунтовых вод (т. е. когда вся жидкость находится в пористой среде) нельзя было бы использовать закон Дарси во всей рассматриваемой области. Разумеется, выполнение этих и других предположений должно контролироваться при изучении данного объекта на основе построенной модели.

При введении дополнительных предположений общая модель упрощается. Так, если по каким-то причинам решение не зависит от времени t (стационарный процесс), то приходим к эллиптическому уравнению

$$\frac{\partial}{\partial x}\left[(H+h)\,\frac{\partial h}{\partial x}\right] + \frac{\partial}{\partial y}\left[(H+h)\,\frac{\partial h}{\partial y}\right] = 0, \tag{12}$$

для решения которого, естественно, не требуется задание функции h в начальный момент. В простейшем случае (12) превращается в *уравнение Лапласа* (см. упр. 3). Если подстилающая поверхность горизонтальна ($H(x,y) = H_0 = \mathrm{const}$), то уравнение Буссинеска становится однородным:

$$\frac{\partial h}{\partial t} = k\,\frac{\partial}{\partial x}\left(h\,\frac{\partial h}{\partial x}\right) + k\,\frac{\partial}{\partial y}\left(h\,\frac{\partial h}{\partial y}\right).$$

При дополнительном предположении об одномерности течения, когда искомое решение зависит лишь от одной пространственной переменной, например, координаты x, приходим к уравнению

$$\frac{\partial h}{\partial t} = k\,\frac{\partial}{\partial x}\left(h\,\frac{\partial h}{\partial x}\right), \tag{13}$$

называемому также *одномерным уравнением типа нелинейной теплопроводности* (см. § 2) или одномерным уравнением изотермической фильтрации. Одномерными являются, например, течения в пластах, сильно вытянутых по одному из направлений, так что изменением величин вдоль поперечного сечения пласта можно пренебречь (если через ограничивающие его в поперечных направлениях поверхности жидкость не протекает). Наконец, самая простая модель течения грунтовых вод дается *уравнением теплопроводности* (или *уравнением диффузии вещества*)

$$\frac{dh}{dt} = kH_0\,\frac{d^2h}{dx^2}, \tag{14}$$

получающимся при условии $h \ll H_0$, т. е. для малых изменений уровня жидкости по сравнению с толщиной пласта.

Последние три уравнения относятся к параболическому типу, причем уравнение (14) линейное и существуют хорошо известные методы получения его общего решения. Разумеется, кроме перечисленных возможны и другие упрощения исходной модели, например двумерное уравнение (13).

Из уравнения Буссинеска относительно нетрудно получить и более сложные модели, когда неверны некоторые из сформулированных

в п. 2 предположений. В частности, во многих случаях грунт неоднороден, т. е. $m = m(x, y)$, $\mu = \mu(x, y)$, и необходимо учитывать поступление жидкости в пласт в результате осадков. Тогда обобщение уравнения Буссинеска имеет вид

$$\frac{m(x,y)}{\rho g}\frac{dh}{dt} = -\frac{\partial}{\partial x}\left[\mu(x,y)(H+h)\frac{\partial h}{\partial x}\right] - \frac{\partial}{\partial y}\left[\mu(x,y)(H+h)\frac{\partial h}{\partial y}\right] + q(x,y,t), \tag{15}$$

где $q(x, y, t)$ характеризует мощность осадков в точке x, y в момент времени t (см. упр. 4).

Итак, применение фундаментального закона сохранения массы позволило получить разнообразные модели рассматриваемых процессов. Различие между моделями определяется типом полученных уравнений (гиперболический, параболический, эллиптический), их пространственно-временными характеристиками (стационарное, нестационарное, одномерное, многомерное), наличием или отсутствием нелинейностей, а также постановкой краевых условий. Таким образом, в зависимости от конкретных свойств объекта и дополнительных предположений, основываясь на одном и том же фундаментальном законе, можно получить совершенно различные математические модели. С другой стороны, как будет не раз показано в дальнейшем, одни и те же математические модели могут, в силу своей универсальности, отвечать объектам совершенно разной природы.

УПРАЖНЕНИЯ

1. Найдите замену переменных, сводящую уравнение (1) к уравнению (3), и покажите, что решение в случае $u = u(t)$ имеет вид (4), где $\xi = x + \int_0^t u(t)\,dt$.

2. Решение уравнения (13) вида $h(x, t) = u(t)\,\theta(x)$ (т. е. в разделяющихся переменных) называется *упорядоченным режимом Буссинеска* в случае, если $u(t) \to 0$, $t \to \infty$. Покажите, что $u(t)$ — степенная функция времени при больших t.

3. Установите, при каких предположениях уравнение (12) сводится к уравнению Лапласа.

4. Пользуясь законом сохранения массы и законом Дарси, получите уравнение (15).

§ 2. Сохранение энергии

Закон сохранения энергии вместе с некоторыми дополнительными предположениями применим для построения моделей распространения тепла в сплошной среде. Сформулируем типичные краевые задачи для уравнений теплопередачи. Обсудим некоторые физические и математические свойства полученных моделей.

1. Предварительные сведения о процессах теплопередачи. Тепловая энергия, или тепло — это энергия хаотического движения атомов или молекул вещества. Обмен теплом между различными участками материала называется *теплопередачей*, а сами материалы, обладающие хорошо выраженным свойством теплопередачи, — *теплопроводными*. К ним относятся, например, металлы, в которых тепловая энергия переносится в основном свободными электронами, некоторые газы и т. д. Процессы передачи тепла рассматриваются в условиях так называемого *локального термодинамического равновесия* (ЛТР). Понятие ЛТР для газов вводится при $\lambda \ll L$, т. е. когда длина свободного пробега частиц вещества много меньше характерных размеров рассматриваемого объекта (*сплошная среда*). ЛТР подразумевает также, что процессы изучаются при временах, больших, чем τ (время между столкновениями частиц), и на размерах, больших, чем λ. Тогда в областях вещества, размеры которых превосходят величину λ (но много меньше величины L), устанавливается равновесие и для них можно ввести средние величины плотности, скорости теплового движения частиц и т. д. Эти локальные величины (разные в разных точках среды) при сформулированных предположениях находятся из равновесного максвелловского распределения частиц (см. § 3 гл. III). К ним относится и *температура* T, определяющая среднюю кинетическую энергию частиц:

$$\frac{mv^2}{2} = \frac{3}{2}kT,$$

где m — масса частицы, v — средняя скорость хаотического движения, k — постоянная Больцмана (в случае так называемого больцмановского газа).

Связанная с хаотическим движением частиц энергия вещества (внутренняя энергия) определяется через температуру с помощью величины *удельной теплоемкости* $c(\rho, T)$, а именно

$$c(\rho,T) = \frac{\partial\varepsilon(\rho,T)}{\partial T}, \qquad c(\rho,T) > 0,$$

где $\rho = mn$ — плотность вещества (n — число частиц в единице объема), $\varepsilon(\rho, T)$ — *внутренняя энергия единицы массы*. Другими словами, теплоемкость — это энергия, которую надо сообщить единице массы вещества, чтобы увеличить его температуру на один градус.

Наиболее простое выражение для теплоемкости получается в случае *идеального газа* (газа, частицы которого взаимодействуют лишь при непосредственном столкновении и, подобно биллиардным шарам, без потери суммарной кинетической энергии). Если в некотором объеме идеального газа содержится N частиц, то их полная внутренняя энергия есть

$$E = N\frac{mv^2}{2} = \frac{3}{2}NkT = \frac{3}{2}M\frac{k}{m}T,$$

где $M = Nm$ — суммарная масса частиц, а удельная внутренняя энергия, или энергия на единицу массы, дается формулой

$$\varepsilon = \frac{E}{M} = \frac{3}{2}\frac{k}{m}T,$$

т. е. теплоемкость идеального газа равна $3k/(2m)$ и не зависит от величин ρ, T.

В общем случае связь между внутренней энергией и температурой более сложная. Например, помимо кинетической энергии движущихся частиц, внутренняя энергия содержит составляющую, связанную с потенциальной энергией их взаимодействия, зависящей от среднего расстояния r между ними. В свою очередь $r \approx (n)^{-1/3} = (\rho/m)^{-1/3}$, где n — число частиц в единице объема, т. е. ε зависит от плотности ρ. Поэтому в теории теплопередачи величины ε (или, что то же самое, c) являются, вообще говоря, функциями от ρ и T. Их конкретный вид определяется свойствами рассматриваемой среды.

2. Вывод закона Фурье из молекулярно-кинетических представлений. Для получения математической модели теплопередачи необходимо, помимо описанных в п. 1 понятий, ввести важное понятие потока тепла. *Потоком тепла* (или тепловой энергии) в данной точке называется количество тепла, переносимое в единицу времени через единичную поверхность, помещенную в данную точку вещества (ср. с понятием потока массы в § 1). Очевидно, что поток тепла — векторная величина (поскольку она в общем случае зависит от ориентации единичной поверхности в пространстве).

Выделим в среде точку с координатами x, y, z и вычислим компоненты потока W тепла по соответствующим осям (величины W_x, W_y, W_z). Расположим площадку единичной величины (штриховая линия на рис. 25) перпендикулярно оси x. Частицы, движущиеся вдоль оси x, пересекают ее справа налево и слева направо с равной вероятностью. Однако если температуры частиц (а, следовательно, и их кинетические энергии) разные по правую и левую стороны площадки, то в единицу времени через нее справа и слева переносятся разные энергии. Разность этих энергий и формирует поток тепла вдоль оси x. Выделим на рис. 25 области, отстоящие на расстояние $\lambda = v\tau$ от площадки справа и слева. Из частиц, находящихся в правой области, примерно 1/6 часть движется налево, так как все шесть направлений (вверх — вниз, вперед — назад, направо — налево) равновероятны. За время τ эта часть частиц с необходимостью пересечет площадку и перенесет энергию, равную

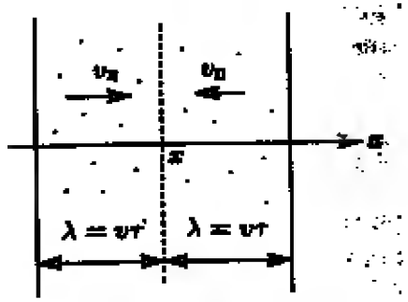


Рис. 25

$$\frac{1}{6} n\lambda \frac{mv_п^2}{2},$$

где $v_п$ — скорость частиц в правой области (величины n, λ считаются в первом приближении равными по обе стороны площадки). Аналогично, частицы из левой области переносят энергию

$$\frac{1}{6} n\lambda \frac{mv_л^2}{2},$$

где $v_л$ — скорость частиц слева от площадки. Разность этих энергий,

отнесенная к единице времени, представляет собой величину

$$W_x = \frac{1}{6} nv \left(\frac{mv_л^2}{2} - \frac{mv_п^2}{2} \right) = \frac{mnv}{6} (\varepsilon_л - \varepsilon_п),$$

где $\varepsilon_л$, $\varepsilon_п$ — внутренняя энергия вещества соответственно слева и справа от площадки, а в качестве v берется средняя между $v_л$ и $v_п$ скорость частиц. В первом приближении величины $\varepsilon_л$, $\varepsilon_п$ можно выразить через величину ε (энергию в точке x, т. е. на площадке) следующим образом:

$$\varepsilon_п = \varepsilon + \lambda \frac{\partial \varepsilon}{\partial x} = \varepsilon + \lambda c \frac{\partial T}{\partial x},$$

$$\varepsilon_л = \varepsilon - \lambda \frac{\partial \varepsilon}{\partial x} = \varepsilon - \lambda c \frac{\partial T}{\partial x}.$$

Подставляя эти формулы в выражение для W_x, получаем

$$W_x = -\varkappa \frac{\partial T}{\partial x}, \tag{1}$$

где $\varkappa = \rho c \lambda v/3$.

Проводя такие же рассуждения для компонент W_y, W_z, приходим к выражениям

$$W_y = -\varkappa \frac{\partial T}{\partial y}, \qquad W_z = -\varkappa \frac{\partial T}{\partial z}. \tag{2}$$

Объединение (1) и (2) дает *закон Фурье*

$$W = -\varkappa \operatorname{grad} T. \tag{3}$$

Величина $\varkappa$ называется *коэффициентом теплопроводности*.

Заметим, что коэффициент теплопроводности зависит в общем случае от плотности и температуры вещества:

$$\varkappa = \frac{\rho c \lambda v}{3} \geqslant 0, \tag{4}$$

поскольку не только теплоемкость c, но и длина свободного пробега λ также может быть функцией от ρ, T. Так, например, в газе, находящемся в обычных условиях, тепло переносится молекулами (молекулярная теплопроводность). Для величины λ в этом случае справедливо $\lambda \sim 1/\rho$, а так как $v \sim \sqrt{T}$, то из (4) имеем $\varkappa_m \sim \sqrt{T}$ (теплоемкость считается постоянной). В плазме (где основную роль в переносе тепла играют электроны) длина пробега электрона зависит от ρ, T, так что $\lambda \sim T^2 \rho^{-1}$, и для величины $\varkappa_e$ справедливо $\varkappa_e \sim T^{5/2}$ (c — постоянная).

Итак, закон Фурье гласит: поток тепла пропорционален градиенту температуры. Так как тепловая энергия непосредственно связана с температурой, то в определенном смысле можно считать, что «поток» температуры пропорционален градиенту самой температуры. Совершенно таким же свойством обладает близкий по сущности процесс

диффузии вещества (закон Фика). Аналогичную интерпретацию можно придать закону Дарси (10) из § 1, хотя движение грунтовых вод по своей природе принципиально отличается от процесса диффузии тепла (и закон Дарси не имеет столь относительно простого теоретического обоснования, как законы Фурье и Фика).

3. Уравнение баланса тепла. Применим закон сохранения энергии для математического описания процесса теплопередачи. Будем при этом считать, что внутренняя энергия вещества изменяется лишь благодаря механизму теплопроводности, т. е. другие виды энергии полагаем несущественными (например, пренебрегаем изменением внутренней энергии за счет химических реакций или за счет работы сил давления, сжимающих некоторый объем газа, и т. д.).

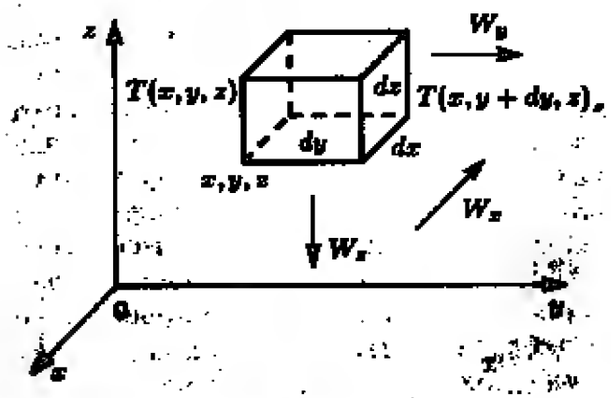


Рис. 26

Выделим в теплопроводной среде элементарный кубик со сторонами dx, dy, dz (рис. 26) и приведем подсчет изменения содержащейся в нем тепловой энергии за малый промежуток времени dt. По сделанным предположениям это изменение может быть вызвано лишь разностью потоков тепла, входящих и выходящих через разные грани кубика. Так, потоки вдоль оси x приводят к уменьшению или увеличению внутренней энергии объема на величину

$$[W_x(x,y,z,t) - W_x(x+dx,y,z,t)]\,dy\,dz\,dt,$$

где $dx\,dy$ — площадь грани, перпендикулярной оси x. В этой формуле считается, что W_x как функция времени не сильно изменяется за промежуток dt, и можно взять ее значение в момент t. Точно таким же образом вычисляются изменения внутренней энергии по осям y, z:

$$[W_y(x,y,z,t) - W_y(x,y+dy,z,t)]\,dx\,dz\,dt,$$

$$[W_z(x,y,z,t) - W_z(x,y,z+dz,t)]\,dx\,dy\,dt.$$

Суммарное изменение энергии $\Delta E = E(t+dt) - E(t)$ есть

$$\Delta E = -\operatorname{div} W\,dx\,dy\,dz\,dt.$$

С другой стороны, величину ΔE можно выразить через изменение температуры объема и через его теплоемкость по формуле

$$\Delta E = (T(t+dt) - T(t))\,c(\rho,t)\,\rho\,dx\,dy\,dz,$$

в которой из-за малости объема берутся некоторые средние по нему значения температуры и плотности.

Приравнивая два последних выражения друг другу и устремляя dt к нулю, получаем *общее уравнение, описывающее распространение*

тепла:

$$C\frac{\partial T}{\partial t} = \operatorname{div}(\varkappa \operatorname{grad} T), \tag{5}$$

имеющее в развернутой форме вид

$$C\frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(\varkappa\frac{\partial T}{\partial x}\right) + \frac{\partial}{\partial y}\left(\varkappa\frac{\partial T}{\partial y}\right) + \frac{\partial}{\partial z}\left(\varkappa\frac{\partial T}{\partial z}\right), \tag{6}$$

где $C = \rho c$.

Уравнение (6) — нестационарное, трехмерное (функция T зависит от времени t и трех пространственных переменных x, y, z) уравнение параболического типа. Оно неоднородное, так как теплоемкость, коэффициент теплопроводности и плотность могут быть, вообще говоря, разными в разных точках вещества, и нелинейное, поскольку функции c и $\varkappa$ могут зависеть от температуры T (т. е. от искомого решения).

При дополнительных предположениях о характере процесса теплопередачи уравнение (6) может упрощаться. Так, если процесс стационарный, т. е. температура не зависит от времени, то (6) превращается в уравнение *эллиптического типа*

$$\frac{\partial}{\partial x}\left(\varkappa\frac{\partial T}{\partial x}\right) + \frac{\partial}{\partial y}\left(\varkappa\frac{\partial T}{\partial y}\right) + \frac{\partial}{\partial z}\left(\varkappa\frac{\partial T}{\partial z}\right), \tag{7}$$

а если функции c, $\varkappa$ не зависят от температуры, то (6) становится *линейным параболическим уравнением*, которое в случае однородной среды ($\varkappa$, c, ρ не зависят от x, y, z) принимает вид

$$\frac{\partial T}{\partial t} = k_0\,\Delta T, \tag{8}$$

где величина $k = \varkappa/c$ называется *коэффициентом температуропроводности*. Для уравнения (8) относительно нетрудно выписать общее решение.

В одномерном случае (температура зависит лишь от t и x) из (6) получаем

$$C\frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(\varkappa\frac{\partial T}{\partial x}\right). \tag{9}$$

Уравнение (9) сводится к уравнению типа нелинейной теплопроводности

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(k(u)\frac{\partial u}{\partial x}\right) \tag{10}$$

при допущении, что $\partial c/\partial x \equiv 0$, $\partial\varkappa/\partial x \equiv 0$ (ср. с уравнением (13) из § 1). Наконец, если $\varkappa = \varkappa_0$, $c = c_0$, где $\varkappa_0$, c_0 — постоянные, то из (10) получается уравнение теплопроводности — простейшее уравнение параболического типа

$$\frac{\partial u}{\partial t} = k_0\frac{\partial^2 u}{\partial x^2}. \tag{11}$$

Как и в случае уравнения Буссинеска, из основного уравнения (6) можно получить различные обобщения, соответствующие более сложным, чем рассмотреные выше, механизмам теплопередачи. Так, для неизотропной среды (т. е. когда коэффициенты теплопроводности разные по разным направлениям) с энерговыделением вместо (6) имеем

$$C\frac{\partial T}{\partial t}=\frac{\partial}{\partial x}\left(\varkappa_x\frac{\partial T}{\partial x}\right)+\frac{\partial}{\partial y}\left(\varkappa_y\frac{\partial T}{\partial y}\right)+\frac{\partial}{\partial z}\left(\varkappa_z\frac{\partial T}{\partial z}\right)+f(x,y,z,t,T), \tag{12}$$

где $\varkappa_x$, $\varkappa_y$, $\varkappa_z$ — коэффициенты в законе Фурье (3) по осям x, y, z, а функция f — мощность выделения (или поглощения) энергии. Неизотропность, например, в случае электронной теплопроводности, может вызываться достаточно сильным магнитным полем, затрудняющим движение переносчиков тепла поперек силовых линий поля, а выделение энергии может быть связано с идущими в веществе химическими реакциями или протеканием электрического тока.

Все полученные в данном пункте уравнения выведены с помощью фундаментального закона сохранения энергии и закона Фурье (ср. с выводом уравнения Буссинеска в § 1). Вместе с заданными функциями c, $\varkappa$, ρ и краевыми условиями они представляют собой замкнутые математические модели процесса теплопередачи.

4. Постановка типичных краевых условий для уравнения теплопроводности. Для простоты будем рассматривать одномерные процессы теплопроводности. Они имеют место, например, в длинном и тонком металлическом стержне (рис. 27), нагреваемом с одного из торцов, при условии, что стержень изотропен, его начальная температура в любом поперечном сечении не зависит от y, z (это же свойство должно соблюдаться и на торцах стержня), а потерями тепла с боковой поверхности можно пренебречь. Будем считать также, что теплоемкость стержня постоянна. Тогда температура зависит только от x и t, и ее распределение вдоль стержня в различные моменты времени описывается уравнением

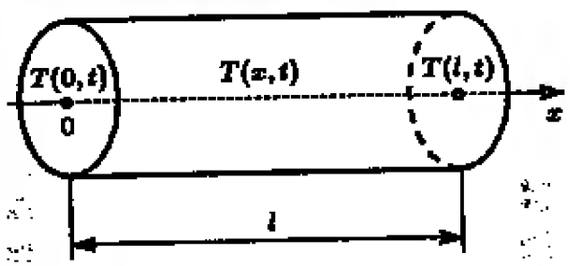


Рис. 27

$$\frac{\partial T}{\partial t}=\frac{\partial}{\partial x}\left(k(T)\frac{\partial T}{\partial x}\right), \tag{13}$$

справедливым при $0<x<l$, $t>0$. Для определения функции $T(x,t)$, т. е. решения, достаточно задать начальную температуру стержня:

$$T(x,t)=T_0(x),\qquad 0\leqslant x\leqslant l, \tag{14}$$

и знать температуру на концах стержня в любой момент времени:

$$T(0,t)=T_1(t),\quad T(l,t)=T_2(t),\qquad t>0. \tag{15}$$

Задача (13)–(15) называется *первой краевой задачей* для параболического уравнения (13) *на отрезке* $x\in[0,l]$. Физически условие (15) соответствует тому, что на концах стержня с помощью каких-то внешних источников тепла поддерживается определенная температура, зависящая, вообще говоря, от времени.

Если же на торцах стержня задаются вместо (15) потоки тепла как функции времени:

$$-k(T(0,t))\left.\frac{\partial T}{\partial x}\right|_{x=0}=W_1(t),\quad k(T(l,t))\left.\frac{\partial T}{\partial x}\right|_{x=l}=W_2(t),\quad t>0, \tag{16}$$

то такая задача называется *второй краевой задачей на отрезке* $[0,l]$. Данная ситуация реализуется, например, когда торцы стержня нагреваются лучами лазерного света известной мощности.

Более сложный (нелинейный) вариант условий на торцах отвечает сильно нагретому и поэтому излучающему энергию стержню, не контактирующему с какими-либо телами. Тогда в единицу времени стержень теряет на своих границах (торцах) энергию, равную $\sigma T^4(0,t)$ и $\sigma T^4(l,t)$ соответственно, и вместо (15) или (16) получаются условия

$$\sigma T^4(0,t)=k(T(0,t))\left.\frac{\partial T}{\partial x}\right|_{x=0},\quad \sigma T^4(l,t)=-k(T(l,t))\left.\frac{\partial T}{\partial x}\right|_{x=l},\quad t>0, \tag{17}$$

где $\sigma>0$.

Возможны также и иные виды краевых условий, соответствующие иным физическим ситуациям. Разумеется, допустимы различные комбинации условий (15)–(17), например, на левом конце известна температура, а на правом поток тепла, и т. д.

Разнообразие постановок краевых условий для уравнений теплопередачи связано и с различными идеализациями исходной задачи (13)–(15). При анализе распространения тепла около одного из торцов длинного стержня в течение сравнительно короткого времени влиянием другого торца можно пренебречь. Вместо (15) достаточно задать лишь одно из условий (для определенности на левом конце):

$$T(0,t)=T_1(t),\qquad t>0, \tag{18}$$

и решать уравнение в области $x>0$ ((13), (14), (18) — *первая краевая задача в полупространстве*).

Обсуждавшаяся уже на примере уравнения Буссинеска (§ 1) задача Коши рассматривается во всем пространстве $-\infty<x<\infty$. Для уравнения (13) задается лишь начальное распределение температуры (14). Такая постановка вполне разумна, когда рассматриваются процессы в центральной части стержня и влияние обоих торцов можно считать несущественным.

Для многомерных уравнений теплопроводности постановка краевых условий по сравнению с одномерным случаем существенно не меняется: на границах области задаются либо температура, либо поток тепла, либо какие-то более сложные их комбинации, а также (в

момент $t = 0$) начальное распределение температуры. Заметим, что в случае стационарного уравнения (7) задаются лишь граничные условия. Краевые условия для уравнения движения грунтовых вод из § 1 вполне аналогичны описанным в этом пункте (при этом аналогами температуры и потока тепла в уравнении Буссинеска служат уровень грунтовых вод и поток массы).

5. Об особенностях моделей теплопередачи. Наиболее простая из всех обсуждавшихся выше задач теплопроводности — задача о стационарном процессе для уравнения (11) на отрезке $[0, l]$:

$$k_0 \frac{\partial^2 T}{\partial x^2} = 0, \qquad T(0) = T_1, \quad T(l) = T_2.$$

Ее решение — линейная функция координаты x:

$$T(x) = \frac{T_2 - T_1}{l} x + T_2, \qquad 0 < x < l. \tag{19}$$

Решение (19) имеет вполне очевидный физический смысл. Действительно, при стационарном процессе потоки тепла, входящие в любое поперечное сечение стержня и выходящие из него, равны (иначе температура в сечении менялась бы). Поэтому поток должен быть постоянен в любой точке x, что по закону Фурье (3) при $\varkappa = \varkappa_0 = \mathrm{const}$ возможно лишь при линейном «профиле» температуры.

Вместе с тем применение закона Фурье приводит к появлению одного не имеющего физического смысла эффекта, характерного для уравнений параболического типа. Поясним его, рассмотрев для уравнения (11), решаемого во всем пространстве $-\infty < x < \infty$, задачу о так называемом *мгновенном точечном источнике тепла*. Требуется найти распределение температуры при всех $t > 0$, $-\infty < x < \infty$, вызванное выделением в момент $t = 0$ в плоскости $x = 0$ некоторого количества тепла Q_0. Начальная температура считается равной нулю: $T(x, 0) = T_0(x) \equiv 0$, $-\infty < x < \infty$. Такая постановка — идеализация реального процесса, справедливая при выполнении соответствующих условий (например, по центру холодного стержня пропускается мощный поперечный импульс электрического тока, действующего очень короткое время и затрагивающего малый участок металла). Решение поставленной таким образом задачи дается формулой

$$T(x, t) = \frac{Q_0}{2\sqrt{\pi k_0 t}} \exp\left(-\frac{x^2}{4k_0 t}\right), \qquad t > 0, \tag{20}$$

что проверяется непосредственной подстановкой в уравнение (11). Симметричная функция (20) в силу известного равенства

$$\int_{-\infty}^{\infty} e^{-y^2}\, dy = \sqrt{\pi}$$

обладает свойством

$$\int_{-\infty}^{\infty} T(x, t)\, dx = Q_0, \qquad t > 0,$$

так что закон сохранения энергии выполняется. В то же время согласно (20) температура в любой точке пространства в любой момент $t > 0$ отлична от нуля. Тем самым модель (11) и многие другие модели теплопередачи описывают процессы с бесконечной скоростью распространения возмущений (температура при $t = 0$ была нулевой для всех x).

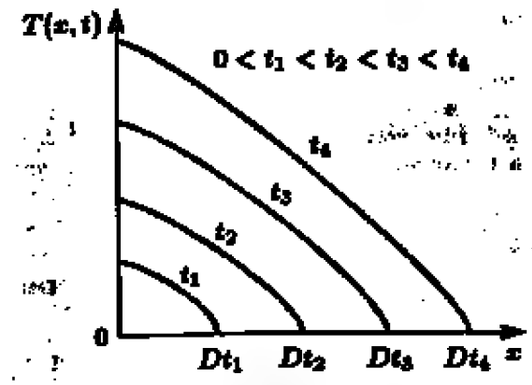


Рис. 28

Этого недостатка лишены (но лишь при определенных условиях) уравнения типа нелинейной теплопроводности (10) (в частности, уравнение (13) § 1). Для модели (10) с $k(T) = k_0 T^\sigma$, $\sigma > 0$ рассмотрим процесс распространения тепла в полупространство $x > 0$ при заданной на границе температуре: $T(0, t) = T_1(t)$. Начальная температура среды считается нулевой: $T(x, 0) = T_0(x) \equiv 0$, $x \geqslant 0$. Частное решение этой задачи, отвечающее граничному закону

$$T_1(t) = \left(\frac{\sigma D^2}{k_0} t\right)^{1/\sigma}, \qquad t > 0,$$

имеет вид бегущей волны (ср. с решением (4) § 1), распространяющейся от границы вглубь вещества не с бесконечной, а с конечной скоростью $D > 0$ (рис. 28):

$$T(x, t) = \begin{cases} \left(\dfrac{\sigma D}{k_0}\right)^{1/\sigma} (Dt - x)^{1/\sigma}, & x \leqslant Dt, \\ 0, & x > Dt, \end{cases} \qquad t > 0. \tag{21}$$

Однако это свойство реализуется лишь при распространении тепла в холодную среду и теряется в случае отличной от нуля начальной температуры вещества (более подробно этот круг вопросов рассматривается в гл. V).

Описанный дефект, связанный с неприменимостью закона Фурье (и закона Дарси в случае уравнения Буссинеска) в окрестности фронта распространения тепловой энергии, не препятствует широкому применению параболических уравнений (из (20) видно, что доля энергии, содержащейся в веществе при достаточно больших значениях x, ничтожна

тожно мала в сравнении с полной энергией Q_0). Они служат хорошим примером универсальности математических моделей, описывая большое количество разнообразных процессов (см. также § 1 гл. IV), имеющих принципиально разную природу.

УПРАЖНЕНИЯ

1. Найдите интегральную замену функции T, с помощью которой уравнение (9) принимает вид (10).

2. Используя те же рассуждения, что и при выводе закона (3) и уравнения (6), получите уравнение (12).

3. Получите решение (20) для задачи о мгновенном точечном источнике тепла, представляя его в виде $T(x,t) = f(t)\varphi(\xi)$, где $\xi = x/\sqrt{4k_0 t}$.

4. Постройте решение (21) для уравнения (10) с $k(T) = k_0 T^\sigma$, $\sigma > 0$, представив температуру в виде $T(x,t) = Af(\xi)$, $A > 0$, $\xi = Dt - x$. Убедитесь в том, что в точке фронта волны $x = Dt$ поток тепла равен нулю.

§ 3. Сохранение числа частиц

Введем некоторые понятия теории теплового излучения, переносимого в среде световыми квантами. Закон сохранения числа квантов используем для получения кинетического уравнения, которому подчиняется функция распределения фотонов. Обсудим некоторые свойства построенной модели лучистого теплообмена в веществе.

1. Основные понятия теории теплового излучения. В веществе, нагретом до достаточно высокой температуры, большую роль играют процессы переноса энергии световыми квантами (фотонами). Распространяясь в среде, рассеиваясь и поглощаясь на атомах и молекулах вещества, а также испускаясь ими, фотоны обеспечивают лучистый теплообмен между различными участками среды. Благодаря именно этому механизму горящий камин нагревает воздух в помещении.

Поле излучения, заполняющее пространство, можно рассматривать как электромагнитное излучение с частотой колебаний ν и длиной волны λ, связанными через скорость света c $(\lambda = c/\nu)$. Если же говорить о поле излучения как о совокупности большого числа частиц — световых квантов, то необходимо ввести понятие энергии кванта $h\nu$ (h — постоянная Планка), движущегося со скоростью c. В отличие от поля температур, характеризуемого координатами x, y, z и временем t, для описания излучения важно знать также его частоту ν (вообще говоря, разную для разных квантов) и направление движения квантов в любой точке пространства в любой момент t.

Проследить траекторию каждого из огромного числа фотонов в веществе попросту невозможно. Поэтому в теории излучения используется статистический вероятностный подход, основанный на введении *функции распределения частиц*. Это важное понятие успешно используется для изучения совокупности большого числа частиц или иных объектов в различных областях знания (см., например, § 3 гл. III).

Функция распределения фотонов $f = f(\nu, \vec{r}, \vec{\Omega}, t)$ зависит от частоты квантов, радиуса-вектора $\vec{r}$ (т. е. от координат x, y, z), направления движения частиц $\vec{\Omega}$ и времени t. Ее смысл состоит в следующем. Рассмотрим в некоторый момент времени t элемент объема $d\vec{r}$ около точки $\vec{r}$ (рис. 29). Тогда величина

$$f(\nu, \vec{r}, \vec{\Omega}, t)\, d\nu\, d\vec{r}\, d\vec{\Omega} \qquad (1)$$

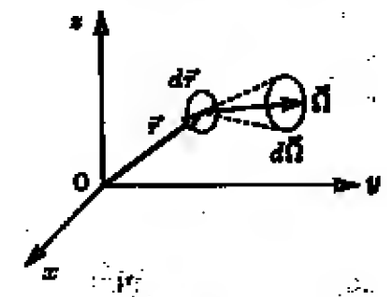


Рис. 29

по определению — это число квантов, находящихся в спектральном интервале $(\nu, \nu + d\nu)$ (т. е. их частота лежит между значениями ν и $\nu + d\nu$), занимающих объем $d\vec{r}$ и имеющих направление движения в диапазоне от $\vec{\Omega}$ до $\vec{\Omega} + d\vec{\Omega}$ ($\vec{\Omega}$ — единичный вектор). Размер объема $d\vec{r}$ предполагается гораздо больше длины волны λ, так что волновые эффекты несущественны.

Функция распределения (1) — одно из исходных понятий теории лучистого теплообмена, с помощью которого вводятся и вычисляются все остальные характеристики, описывающие этот процесс. Величина I_ν, определяемая по формуле

$$I_\nu(\vec{r}, \vec{\Omega}, t) = h\nu c f(\nu, \vec{r}, \vec{\Omega}, t), \qquad (2)$$

называется *спектральной интенсивностью излучения*. Она представляет собой количество лучистой энергии в спектральном интервале от ν до $\nu + d\nu$, переносимое фотонами за единицу времени через единичную площадку, помещенную в точке $\vec{r}$ и перпендикулярную к направлениям их полета (которые лежат в диапазоне углов от $\vec{\Omega}$ до $\vec{\Omega} + d\vec{\Omega}$; рис. 30).

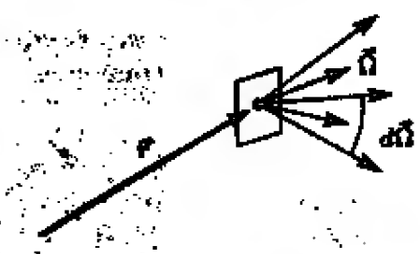


Рис. 30

Действительно, так как энергия кванта равна $h\nu$, а общее число квантов с частотой от ν до $\nu + d\nu$ и с направлением полета от $\vec{\Omega}$ до $\vec{\Omega} + d\vec{\Omega}$ в единице объема равно $f\, d\nu\, d\vec{\Omega}$, то переносимая ими за 1 с через расположенную перпендикулярно полету площадку в 1 см^2 энергия равна $h\nu c f\, d\nu\, d\vec{\Omega}$, что согласуется с определением (2).

Спектральная плотность излучения

$$U_\nu(\vec{r}, t) = h\nu \int\limits_{4\pi} f\, d\vec{\Omega} = \frac{1}{c} \int\limits_{4\pi} I_\nu\, d\vec{\Omega} \qquad (3)$$

представляет собой количество лучистой энергии квантов, содержащихся в 1 см^3 пространства в точке $\vec{r}$ в момент t в единичном интервале частот и имеющих частоту ν.

Еще одной важной характеристикой служит *спектральный поток излучения* $\vec{S}_\nu$. Фотоны, пересекающие единичную площадку с направлением нормали $\vec{n}$, переносят через нее энергию (в 1 с в интервале от ν

до $\nu + d\nu$) равную $h\nu c \int\limits_{2\pi} f \cos\theta \, d\vec{\Omega}$ (рис. 31). Аналогично вычисляется энергия, распространяющаяся через площадку справа налево, но интегрирование ведется по левой полусфере. Их разность и дает величину S_ν:

Рис. 31

$$S_\nu(\vec{r}, t, \vec{n}) = h\nu c \int\limits_{4\pi} f \cos\theta \, d\vec{\Omega}, \tag{4}$$

где θ — угол между направлением движения квантов и нормалью. Величина S_ν — проекция вектора $\vec{S}_\nu$ на нормаль $\vec{n}$, а сам вектор есть

$$\vec{S}_\nu = \int\limits_{4\pi} I_\nu \vec{\Omega} \, d\vec{\Omega}. \tag{5}$$

Заметим, что при изотропном (не зависящем от направления $\vec{\Omega}$) излучении спектральная плотность, как следует из (3), равна

$$U_\nu = 4\pi h\nu f,$$

а из (5) видно, что поток $\vec{S}_\nu$ равен нулю в любой точке пространства.

Полные интенсивность, плотность и поток излучения можно получить из спектральных характеристик интегрированием по всему спектру частот ν.

2. Уравнение баланса числа фотонов в среде. Выведем уравнение, описывающее перенос излучения в среде, пользуясь законом сохранения числа частиц и следующими предположениями:

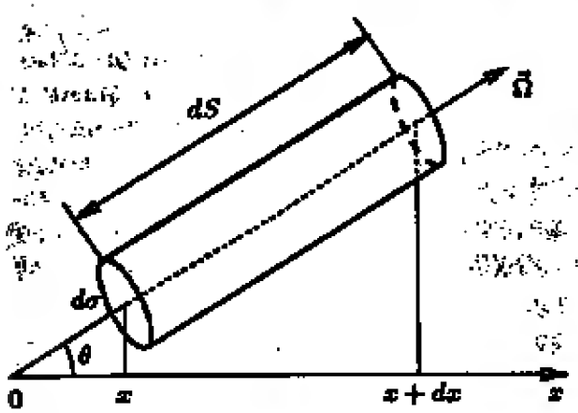


Рис. 32

1) процесс распространения квантов одномерный, т. е. $f = f(\nu, x, \vec{\Omega}, t)$;

2) рассеянием квантов света на атомах или молекулах (т. е. изменением их направления) можно пренебречь;

3) известен характер поглощения и испускания света атомами и молекулами вещества;

4) фотоны самопроизвольно не исчезают и не появляются.

Рассмотрим баланс частиц в элементарном цилиндре, имеющем ось в направлении $\vec{\Omega}$, длину $ds = dx/\cos\theta$ и основание $d\sigma$ (рис. 32), где θ — угол между осью x и вектором $\vec{\Omega}$. Будем интересоваться излучением частоты ν в единичном интервале частот, распространяющимся внутри единичного телесного угла в направлении $\vec{\Omega}$.

В соответствии с (1) (см. также определение (2)) за время dt в левое основание цилиндра входит число частиц, равное

$$cf(\nu, x, \vec{\Omega}, t) \, d\sigma \, dt.$$

За то же время из его правого основания выходит число частиц

$$(cf(\nu, x, \vec{\Omega}, t) + c\,df) \, d\sigma \, dt,$$

где величина df описывает приращение функции f при переходе от одного основания к другому. Поскольку $f = f(\nu, \vec{\Omega}, x, t)$, то эту величину можно представить в виде

$$df = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial s} ds, \qquad ds = \frac{dx}{\cos\theta},$$

где первое слагаемое отвечает ее приращению по времени за промежуток dt, а второе — приращению по координате s.

Учитывая, что скорость фотонов равна c и $dt = ds/c$, получаем

$$df = \left(\frac{1}{c}\frac{\partial f}{\partial t} + \frac{\partial f}{\partial s}\right) ds.$$

Итак, число фотонов в цилиндре за время dt изменилось на величину

$$-\left(\frac{\partial f}{\partial t} + c\frac{\partial f}{\partial s}\right) ds \, d\sigma \, dt = -\left(\frac{\partial f}{\partial t} + c\cos\theta\frac{\partial f}{\partial x}\right) ds \, d\sigma \, dt. \tag{6}$$

Напомним, что через боковую поверхность цилиндра фотоны, имеющие направление полета $\vec{\Omega}$ и не претерпевающие рассеяния, не пролетают.

Таким образом, изменение числа квантов в объеме цилиндра может вызываться лишь их поглощением или испусканием атомами и молекулами вещества, находящегося внутри цилиндра. Для вычисления этой величины вводится понятие *равновесного излучения*, когда число квантов, поглощенных веществом, равно числу испущенных частиц (излучение и вещество находятся в равновесии) в любой момент времени. Равновесная функция распределения $f_р$ есть (*закон Планка*)

$$f_р = \frac{2\nu^2}{c^3} \exp\left(1 - \frac{h\nu}{kT}\right), \tag{7}$$

где T — температура вещества (считается, что среда находится в условиях локального термодинамического равновесия, и в любой ее точке можно ввести такие характеристики, как температура, внутренняя энергия и т. д.).

В отсутствие равновесия между излучением и веществом интенсивность поглощения (испускания) фотонов пропорциональна разности между $f_р$ и f, т. е. величине

$$\varkappa_\nu c(f - f_р),$$

где $\varkappa'_\nu = \varkappa_\nu(1 - \exp(h\nu/kT))$, а $\varkappa_\nu$ — коэффициент поглощения, определяемый состоянием среды и ее свойствами. Изменение числа квантов в объеме цилиндра за время dt равно

$$\varkappa'_\nu (f - f_p)\, d\sigma\, ds\, dt. \tag{8}$$

Приравнивая (6) и (8), получаем для функции распределения *кинетическое уравнение*, описывающее перенос излучения в среде:

$$\frac{\partial f}{\partial t} + c\cos\theta\,\frac{\partial f}{\partial x} = \varkappa'_\nu (f_p - f), \tag{9}$$

где f_p задается формулой (7). Уравнение (9) вместе с функциями f_p, $\varkappa'_\nu$ и краевыми условиями представляет собой замкнутую модель распространения лучистой энергии при сделанных выше предположениях.

3. Некоторые свойства уравнения переноса излучения. Полученное на основании закона сохранения числа частиц нестационарное одномерное неоднородное гиперболическое уравнение (9) может быть обосновано также и с помощью закона сохранения энергии. Действительно, в цилиндре рассматривался баланс частиц, имеющих одинаковую частоту ν и, следовательно, одинаковую энергию $h\nu$. Учитывая это, (9) легко переписать как уравнение относительно спектральной интенсивности излучения $I_\nu = h\nu c f$:

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \cos\theta\,\frac{\partial I_\nu}{\partial x} = \varkappa_\nu (I_{\nu p} - I_\nu), \tag{10}$$

которое эквивалентно (9), но имеет более непосредственный физический смысл.

При интегрировании (10) по телесному углу $\vec{\Omega}$ (т. е. по всем направлениям полета квантов) получаем уравнение, связывающее плотность излучения (3) и его поток (4):

$$\frac{\partial U_\nu}{\partial t} + \frac{\partial S_\nu}{\partial x} = c\varkappa_\nu (U_{\nu p} - U_\nu). \tag{11}$$

Это уравнение можно трактовать как уравнение неразрывности для излучения данной частоты, выражающее закон сохранения излучения и вполне аналогичное уравнению (7) § 1 в теории движения грунтовых вод и уравнению (5) § 2 в теории теплопроводности. Наиболее очевидна эта аналогия в трехмерном случае, когда уравнения (10) и (11) принимают вид

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \vec{\Omega}\nabla I_\nu = \varkappa_\nu (I_{\nu p} - I_\nu), \tag{12}$$

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \operatorname{div}\vec{S}_\nu = c\varkappa_\nu (U_{\nu p} - U_\nu). \tag{13}$$

Хотя уравнения (9)–(13) линейные, нельзя, вообще говоря, утверждать, что модели лучистого теплообмена проще нелинейных моделей, рассмотренных в § 1, 2. Ведь, решая (9)–(13), можно получать каждый

раз лишь спектральные (т. е. для данной частоты ν) характеристики излучения, распространяющегося в заданном направлении $\vec{\Omega}$. Для полной картины необходимо найти нужные величины для всех значений ν, $\vec{\Omega}$ (или какие-то интегралы от них), что является гораздо более трудной задачей. К тому же в более сложных ситуациях (наличие рассеивания фотонов и т. д.) сами модели (9)–(13) могут значительно усложняться.

Наиболее простая модель переноса излучения получается из (10), если поглощением и испусканием квантов можно пренебречь и рассматривать случай, когда все частицы движутся в одном направлении. Тогда для любых значений ν можно положить $\cos\theta = 1$ и прийти к уравнению

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \frac{\partial I_\nu}{\partial x} = 0,$$

полностью идентичному уравнению (3) § 1 для потока невзаимодействующих материальных частиц.

Если интенсивность излучения не зависит от времени, то (10) превращается в неоднородное линейное дифференциальное уравнение

$$\cos\theta \cdot \frac{dI_\nu}{dx} + \varkappa_\nu I_\nu = \varkappa_\nu I_{\nu p}, \tag{14}$$

общее решение которого имеет вид

$$I_\nu(x) = \int_{x_0}^{x} \varkappa_\nu I_{\nu p} e^{-\varkappa(x')}\, dx' + I_{\nu_0} e^{-\varkappa(x_0)}. \tag{15}$$

Здесь $\varkappa(x') = \int_{x'}^{x} \varkappa_\nu\, dx''$, $\varkappa(x_0) = \int_{x_0}^{x} \varkappa_\nu\, dx''$ (для простоты в (15) положено $\cos\theta = 1$), I_{ν_0} — постоянная интегрирования.

Не останавливаясь подробно на физическом смысле решения (15), поясним, что первый член обязан своим происхождением излучению, возникшему в веществе на отрезке от x_0 до x (и ослабленному поглощением). Второе слагаемое представляет собой излучение от каких-то внешних источников, входящих в вещество на его границе x_0 (и также ослабленное поглощением по мере распространения по среде).

Если $I_{\nu p}$ и $\varkappa_\nu$ — известные функции координаты x (для этого должны быть известны температура и плотность вещества вдоль траектории частиц), то решение уравнения (14) сводится к квадратуре.

В противоположном рассмотренному случае пространственно однородного поля излучения из (10) получаем

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} = \varkappa_\nu (I_{\nu p} - I_\nu). \tag{16}$$

Процесс, описываемый уравнением (16), соответствует ситуации, когда в неограниченной и первоначально холодной среде (т. е. в момент $t = 0$ излучения нет) с постоянной плотностью происходит быстрый нагрев вещества до некоторой температуры T, которая затем

поддерживается неизменной во времени. Поскольку потерь излучения с границ нет, то пространственные градиенты T равны нулю и $\varkappa_\nu$, $I_{\nu p}$ не зависят от x, y, z. Возникшее в результате нагрева излучение также имеет нулевой градиент (т. е. $I_\nu = I_\nu(t)$) и, обмениваясь энергией с веществом, стремится с течением времени к равновесному значению по экспоненциальному закону.

УПРАЖНЕНИЯ

1. Проверьте правильность выражения (5).
2. Повторяя рассуждения п. 2, получите трехмерное уравнение (9) или (10) в случае, когда функция распределения зависит от x, y, z.
3. Используя определения (3), (4), выведите из (12) уравнение (13).
4. Получите решение (15) уравнения (14) и конкретизируйте его в случае постоянных $\varkappa_\nu$, $I_{\nu p}$.
5. Убедитесь в том, что решение уравнения «насыщения» (16) (ср. с уравнением (12) § 1 гл. I) имеет экспоненциальный вид, и найдите показатель экспоненты.

§ 4. Совместное применение нескольких фундаментальных законов

Законы сохранения массы, импульса, энергии используем для построения математической модели, описывающей течение сжимаемого газа. Обсудим отличия полученной модели от моделей, рассматриваемых в § 1–3, а также некоторые следующие из нее свойства газодинамических движений.

1. Предварительные понятия газовой динамики. Заметное изменение плотностей жидкостей и твердых тел может достигаться лишь при огромных давлениях в десятки и сотни тысяч атмосфер и выше. Газообразные среды гораздо легче подвергаются сжатию: при перепаде давления в одну атмосферу плотность газа, первоначально находившегося при атмосферном давлении, уменьшается или увеличивается на величину, сопоставимую с начальной его плотностью.

В газовой динамике, изучающей движение сжимаемых сред под действием каких-либо внешних сил или сил давления самого вещества, считается выполненным неравенство $\lambda \ll L$, где λ — длина свободного пробега, L — характерные размеры области рассматриваемого течения (сплошная среда). Считается также выполненной гипотеза о ЛТР (см. п. 1 § 2). В условиях ЛТР сжимаемую среду можно рассматривать как совокупность большого числа жидких частиц с размерами, много большими λ, но много меньшими, чем L. Для каждой такой частицы, связанной с небольшой фиксированной массой среды, вводятся характеризующие ее средние величины — плотность ρ, давление p, температура T, внутренняя энергия ε и т. д., а также скорость $\vec{v}$ ее макроскопического движения как единого целого. Все эти величины в общем случае зависят от трех пространственных переменных x, y, z и времени t.

В дальнейшем будем также предполагать отсутствие в среде процессов теплопередачи, вязкого трения, источников и стоков энергии,

например, излучения, и, кроме того, отсутствие внешних объемных сил и источников (стоков) массы в веществе.

2. Уравнение неразрывности для сжимаемого газа. Применим рассуждения, аналогичные тем, что использовались для вывода уравнений неразрывности (7) § 1, (5) § 2 для течения грунтовых вод и процесса теплопередачи. Рассмотрим в некоторой области пространства, занятой движущимся газом, элементарный кубик со сторонами dx, dy, dz и подсчитаем в нем баланс массы за время dt (рис. 33). Здесь v_x, v_y, v_z — компоненты скорости по соответствующим осям.

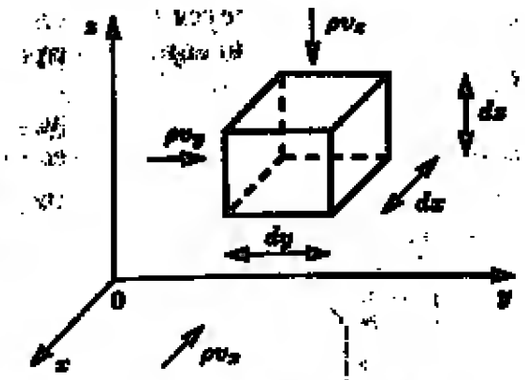


Рис. 33

По оси x через грань с координатой x в кубик за время dt поступает масса газа, равная

$$\rho v_x \, dy \, dz \, dt,$$

поскольку величина ρv_x не что иное, как поток массы по направлению оси x. За то же самое время из грани с координатой $x + dx$ вытекает масса

$$[\rho v_x + d(\rho v_x)] \, dy \, dz \, dt,$$

где через $d(\rho v_x)$ обозначено приращение потока массы при переходе от координаты x к координате $x + dx$. Суммируя оба последних выражения и учитывая, что

$$d(\rho v_x) = \frac{\partial}{\partial x}(\rho v_x)\, dx,$$

получаем величину изменения массы в кубике за время dt благодаря движению газа вдоль оси x:

$$dm_x = -\frac{\partial}{\partial x}(\rho v_x)\, dx\, dy\, dz\, dt. \tag{1}$$

Точно таким же образом находим изменения массы за счет движения по осям y, z:

$$dm_y = -\frac{\partial}{\partial y}(\rho v_y)\, dx\, dy\, dz\, dt,$$
$$dm_z = -\frac{\partial}{\partial z}(\rho v_z)\, dx\, dy\, dz\, dt. \tag{2}$$

В фиксированном объеме кубика изменение находящейся в нем массы газа выражается также через изменение его плотности со временем:

$$dm = \frac{\partial \rho}{\partial t}\, dt\, dx\, dy\, dz. \tag{3}$$

Суммируя dm_x, dm_y, dm_z и приравнивая результат к dm, получаем из (1)–(3) искомое уравнение неразрывности

$$\frac{\partial \rho}{\partial t} + \operatorname{div} \rho \vec{v} = 0, \tag{4}$$

выражающее закон сохранения массы вещества применительно к движению сжимаемого газа. По своей форме и смыслу (скорость изменения величины определяется дивергенцией потока этой величины) оно вполне аналогично уравнению неразрывности (7) § 1 и уравнениям (5) § 2, (11) § 3.

Однако аналогия с течением грунтовых вод на этом заканчивается. При свободном движении газа его динамика определяется лишь силами давления самого газа, в отличие от движения жидкости, испытывающей сопротивление частиц грунта.

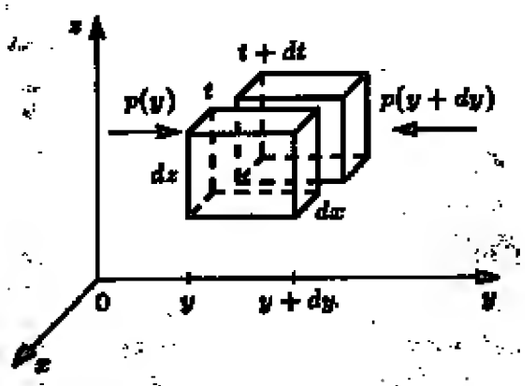


Рис. 34

3. Уравнения движения газа. Для их получения применим второй закон Ньютона к элементарной жидкой частице, имеющей в некоторый момент t форму кубика с гранями dx, dy, dz (рис. 34). Жидкая частица — это перемещающийся в пространстве и меняющий свою форму объем, содержащий в разные моменты времени t одни и те же атомы и молекулы газа. Тем самым его масса dm постоянна. Для простоты вывода будем считать, что за короткое время dt кубик не меняет своей формы и смещается по всем направлениям на расстояние, много меньшее его размеров.

Определим сначала силу, действующую на кубик, например в направлении оси y. Она, очевидно, равна разности давлений на левой и правой гранях, умноженной на их площади (иных сил по предположению нет):

$$F_y = [p(x, y, z, t) - p(x, y + dy, z, t)]\, dx\, dz.$$

Сила F_y равна ускорению жидкой частицы в направлении y, умноженному на ее массу $dm = \rho\, dx\, dy\, dz$:

$$F_y = \frac{dv_y}{dt} \rho\, dx\, dy\, dz. \tag{5}$$

Заменяя в первом выражении для F_y разность давлений через производную от давления по y и приравнивая его к (5), приходим к уравнению, описывающему движение газа вдоль оси y:

$$\rho \frac{dv_y}{dt} = -\frac{\partial p}{\partial y}. \tag{6}$$

Точно так же получаем уравнения движения по направлениям x, z:

$$\rho \frac{dv_x}{dt} = -\frac{\partial p}{\partial x}, \tag{7}$$

$$\rho \frac{dv_z}{dt} = -\frac{\partial p}{\partial z}, \tag{8}$$

имеющие, как и (6), очевидный физический смысл. В векторной форме уравнения (6)–(8) имеют вид

$$\rho \frac{d\vec{v}}{dt} = -\operatorname{grad} p. \tag{9}$$

Поясним, что в (6)–(9) через df/dt обозначена полная (субстанциональная, т. е. связанная с фиксированными частицами газа) производная по времени какой-либо величины, характеризующей данную неизменную массу газа.

Раскрыв df/dt через частные производные по x и t в соответствии с правилом $df/dt = \partial f/\partial t + (\vec{v}\operatorname{grad}) f$, придем к *уравнениям движения Эйлера*

$$\frac{\partial \vec{v}}{\partial t} + (\vec{v}\operatorname{grad})\,\vec{v} = -\frac{1}{\rho}\operatorname{grad} p. \tag{10}$$

Будучи записаны покоординатно, они принимают вид

$$\frac{\partial v_x}{\partial t} + v_x \frac{\partial v_x}{\partial x} + v_y \frac{\partial v_x}{\partial y} + v_z \frac{\partial v_x}{\partial z} = -\frac{1}{\rho}\frac{\partial p}{\partial x}, \tag{11}$$

$$\frac{\partial v_y}{\partial t} + v_x \frac{\partial v_y}{\partial x} + v_y \frac{\partial v_y}{\partial y} + v_z \frac{\partial v_y}{\partial z} = -\frac{1}{\rho}\frac{\partial p}{\partial y}, \tag{12}$$

$$\frac{\partial v_z}{\partial t} + v_x \frac{\partial v_z}{\partial x} + v_y \frac{\partial v_z}{\partial y} + v_z \frac{\partial v_z}{\partial z} = -\frac{1}{\rho}\frac{\partial p}{\partial z}. \tag{13}$$

В отличие от течения грунтовых вод, градиенты давления в уравнениях движения газа (6)–(13) определяют компоненты ускорения вещества, а не компоненты его скорости (ср. с законом Дарси (9) § 1).

Уравнения (4), (11)–(13) содержат пять неизвестных величин — ρ, p, v_x, v_y, v_z. Для их замыкания естественнее всего использовать закон сохранения энергии.

4. Уравнение энергии. Для его получения используем ту же упрощенную схему, что и в п. 3: будем рассматривать изменение внутренней энергии фиксированной массы газа dm за короткий промежуток времени dt. Так как по сделанным допущениям в веществе отсутствует теплопроводность, вязкость и источники (стоки) энергии, то это изменение вызывается лишь работой сил давления на гранях кубика при его сжатии или расширении.

Работа давления, связанная с движением граней объема вдоль оси x, очевидно, равна

$$dA_x = p\,(v_x(x) - v_x(x + dx))\, dt\, dy\, dz,$$

где слагаемые в скобках можно, отбрасывая члены второго порядка малости, переписать через производную $\partial v_x/\partial x$ и получить

$$dA_x = -p\frac{\partial v_x}{\partial x}\, dx\, dy\, dz\, dt.$$

Здесь p — среднее давление в элементарном объеме. Аналогично,

$$dA_y = -p\frac{\partial v_y}{\partial y}\,dx\,dy\,dz\,dt,$$

$$dA_z = -p\frac{\partial v_z}{\partial z}\,dx\,dy\,dz\,dt.$$

Полная работа, совершенная над газом за время dt, есть

$$dA = dA_x + dA_y + dA_z = -p\operatorname{div}\vec{v}\,dx\,dy\,dz\,dt.$$

Она равна изменению внутренней энергии объема, т. е.

$$dA = \rho d\varepsilon\,dx\,dy\,dz,$$

ε — удельная внутренняя энергия. Приравняв оба выражения для dA и устранив к нулю dt, окончательно получим

$$\rho\frac{d\varepsilon}{dt} + p\operatorname{div}\vec{v} = 0, \tag{14}$$

где $d\varepsilon/dt$ — полная (субстанциональная) производная внутренней энергии по времени.

Заметим, что с помощью уравнений неразрывности и движения уравнение (14) приводится, подобно (4), к *дивергентному виду*

$$\frac{\partial}{\partial t}\left(\rho\varepsilon + \frac{\rho v^2}{2}\right) = -\operatorname{div}\left[\rho\vec{v}\left(\varepsilon + \frac{v^2}{2}\right) + p\vec{v}\right]. \tag{15}$$

Слева в (15) стоит производная от полной (внутренней и кинетической) энергии газа в данной точке пространства.

Так как термодинамические свойства вещества предполагаются известными, то ε — известная функция уже введенных величин p и ρ, и уравнение (14) либо (15) дает недостающую связь для определения искомых газодинамических величин.

5. Уравнения газовой динамики в лагранжевых координатах. В полученных моделях течение газа характеризуется зависимостью величин от декартовых координат x, y, z и времени t. Этот способ описания (*эйлеров подход*) трактует движение среды с точки зрения неподвижного стороннего наблюдателя и удобен, например, при изучении обтекания газом моделей летательных аппаратов в аэродинамических трубах. В *лагранжевом подходе* координата связывается не с определенной точкой пространства, а с определенной, фиксированной частицей вещества — жидкой частицей. Лагранжевы координаты удобны, например, при анализе некоторых внутренних процессов, протекающих в частице, скажем, химических реакций, скорость которых определяется не пространственным положением частицы, а ее температурой и плотностью. Лагранжев подход может быть полезен и по другим причинам: он, например, неявно использовался в пп. 3, 4 для упрощенного вывода уравнений движения и энергии (рассматривалась жидкая частица в виде кубика).

Особенно наглядную и простую форму газодинамические уравнения в лагранжевых координатах имеют в одномерном случае. Действительно, взяв в направлении оси x столб единичного сечения, левая граница которого подвижна и связана с фиксированными частицами вещества, можно ввести лагранжеву координату по правилу

$$m(x) = \int_{x_0}^{x}\rho(x,t)\,dx, \qquad dm = \rho\,dx, \tag{16}$$

где $x_0(t)$ — координата частиц столба на его левом конце, $x(t)$ — текущая координата (в качестве частицы с координатой $x_0(t)$ можно выбрать частицу около твердой стенки, ограничивающей газ, или у границы с пустотой, если таковые имеются). Величина $m(x)$ — масса столба между точками x_0, x.

Связь (16) дает наиболее естественный и простой способ введения лагранжевой переменной (называемой в этом случае *массовой координатой*). Теперь все газодинамические величины трактуются как зависящие не от x, t, а от m, t. Получим для них уравнения движения из одномерных уравнений в эйлеровой форме

$$\begin{aligned}
&\frac{\partial\rho}{\partial t} + v\frac{\partial\rho}{\partial x} = -\rho\frac{\partial v}{\partial x},\\
&\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x} = -\frac{1}{\rho}\frac{\partial p}{\partial x},\\
&\frac{\partial\varepsilon}{\partial t} + v\frac{\partial\varepsilon}{\partial x} = -\frac{p}{\rho}\frac{\partial v}{\partial x}.
\end{aligned} \tag{17}$$

Выражения, стоящие в (17) слева, это уже вводившиеся субстанциональные производные, описывающие изменение со временем величин, относящихся к фиксированной массовой координате m. Тогда, используя (16), из (17) получаем

$$\frac{\partial}{\partial t}\frac{1}{\rho} = \frac{\partial v}{\partial m}, \tag{18}$$

$$\frac{\partial v}{\partial t} = -\frac{\partial p}{\partial m}, \tag{19}$$

$$\frac{\partial\varepsilon}{\partial t} = -p\frac{\partial v}{\partial m}. \tag{20}$$

Здесь $\partial/\partial t$ — субстанциональная производная по времени от соответствующих величин. Физическая трактовка уравнений движения (19) и энергии (20) та же, что и в эйлеровых координатах (в отличие от уравнения неразрывности (18)). Последнее представляет собой очевидное свойство: объем (а с ним и плотность) фиксированной жидкой частицы изменяется со временем благодаря разности скоростей на ее границах.

С помощью (19) уравнение энергии можно записать в дивергентном виде

$$\frac{\partial}{\partial t}\left(\varepsilon+\frac{v^2}{2}\right)=-\frac{\partial}{\partial m}(pv), \tag{21}$$

где слева стоит производная по времени относительно полной энергии частицы.

Если решение лагранжевых уравнений найдено, в частности, найден удельный объем $V(m,t)=1/\rho(m,t)$, то зависимость газодинамических функций от эйлеровой координаты находится из квадратуры (см. (16))

$$dx=V(m,t)\,dm,\qquad x(m,t)=\int\limits_0^m V(m,t)\,dm+x_0(t).$$

С помощью таких же рассуждений нетрудно получить лагранжевы уравнения в случае цилиндрической и сферической симметрии, когда газодинамические величины зависят лишь от одной пространственной координаты r (r — расстояние от оси или от центра симметрии) и от времени t. Они также имеют весьма простой и наглядный вид, чего нельзя сказать о двумерном и трехмерном случае.

Несмотря на разные формы записи, эйлеровы и лагранжевы уравнения газовой динамики обладают, естественно, аналогичными свойствами, являясь нелинейными гиперболическими уравнениями в частных производных (нестационарными и в общем случае многомерными). На их основе получаются более сложные модели движения сжимаемых сред, включающие дополнительные физические процессы. Так, для газа, обладающего теплопроводностью, участвующей в процессе переноса энергии, уравнения неразрывности (18) и движения (19) остаются в силе, а уравнение энергии приобретает вид

$$\frac{\partial\varepsilon}{\partial t}=-p\frac{\partial v}{\partial m}-\frac{\partial W}{\partial m}, \tag{22}$$

где $W=-\varkappa\rho\,\partial T/\partial m$ — поток тепла, $\varkappa=\varkappa(\rho,T)$ — коэффициент теплопроводности. Внутренняя энергия жидкой частицы такого газа изменяется не только за счет работы сил давления, но и из-за наличия теплопередачи. Более простые, чем (18)–(20), модели газовой динамики получаются при соответствующих дополнительных предположениях (см. п. 7).

6. Краевые условия для уравнений газовой динамики. Наиболее наглядна их постановка в случае одномерного течения газа, описываемого уравнениями (18)–(20). Рассмотрим такое течение в трубе, внутри которой помещен газ, ограниченный справа и слева непроницаемыми твердыми стенками — поршнями (рис. 35). Частицам, находящимся у левой стенки, припишем координату $m=0$; тогда координата частиц у правой стенки равна $m=M$, где M — полная масса газа между поршнями в столбе единичного сечения. Внутренние

частицы имеют координату $0<m<M$. Уравнения (18)–(20) рассматриваются в области $0<m<M$ и при $t>0$.

Для определения течения при всех $0<m<M$ и $t>0$ необходимо задать:

1) начальные условия, т. е. состояние газа в момент $t=0$,

$$v(m,0)=v_0(m),\quad p(m,0)=p_0(m),\quad \rho(m,0)=\rho_0(m), \tag{23}$$
$$0\leqslant m\leqslant M;$$

в (23) вместо p или ρ можно, пользуясь уравнениями состояния среды, задавать начальную температуру $T(m,0)=T_0(m)$;

2) граничные условия, т. е. зависимость от времени газодинамических величин на границах $m=0$, $m=M$, например, закон изменения давления

$$p(0,t)=p_1(t),\quad p(M,t)=p_2(t),\qquad t>0, \tag{24}$$

или закон изменения скорости поршней (т. е. их траекторию в эйлеровых координатах, так как $\partial x/\partial t=v$)

$$v(0,t)=v_1(t),\quad v(M,t)=v_2(t),\qquad t>0. \tag{25}$$

Знание краевых условий (23) и (24) (либо (25)) полностью определяет однозначное решение рассматриваемой *задачи о поршне*.

В качестве граничных условий допускаются также различные комбинации (24) и (25), когда на левой границе поддерживается известное давление $p_1(t)$, а на правой задается известная скорость $v_2(t)$ (или наоборот, как на рис. 35).

Если интересоваться лишь течением в окрестности одного из поршней, считая влияние второго поршня несущественным, то достаточно задать лишь одно из условий (24) (или (25)), например, при $m=0$, и условие (23) при всех $m>0$ (газ занимает полупространство, ограниченное слева поршнем).

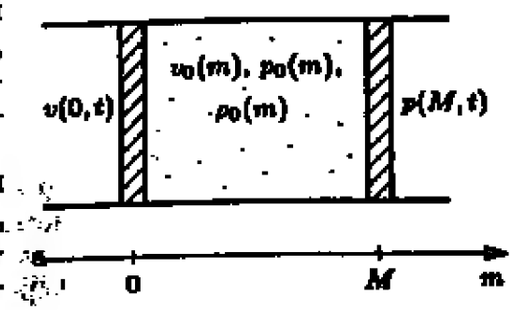


Рис. 35

Наконец, граничные условия (24) или (25) вообще не задаются, когда влиянием границ на движение газа в центральной области можно пренебречь (рассматривая процесс при достаточно малых временах). Тогда из исходной задачи для уравнений (18)–(20) получается задача Коши — задача об эволюции со временем некоторого первоначального распределения газодинамических величин, заданного в неограниченном пространстве. При этом начальные данные (23) определены для всех $-\infty<m<\infty$, а уравнения решаются в области $-\infty<m<\infty$, $t>0$.

Важный класс краевых условий — условия на границе с вакуумом. Пусть сильно сжатый газ, имеющий высокое давление, начинает истекать в сравнительно разреженную среду, находящуюся при низком

давления. Идеализируя этот процесс, можно считать давление и плотность в пространстве, в которое истекает газ, равными нулю, т. е. задавать условие $p_1(t) = 0$, $t > 0$, либо $p_2(t) = 0$, $t > 0$ (либо $p_1(t) = p_2(t) = 0$, $t > 0$), в зависимости от конкретной постановки задачи.

7. Некоторые особенности моделей газовой динамики. Для пояснения этих особенностей предварительно упростим уравнения (18)–(20), используя два обстоятельства. Первое из них — отсутствие в среде (по предположению) потерь энергии за счет теплопроводности, вязкости, излучения, внешних источников и стоков энергии и т. д. С термодинамической точки зрения это означает, что процесс адиабатический и энтропия S каждой фиксированной жидкой частицы со временем не меняется. Тогда уравнение энергии (20) можно переписать в эквивалентной форме

$$\frac{\partial S}{\partial t} = 0. \tag{26}$$

В этом нетрудно убедиться, также чисто формально применяя второе начало термодинамики

$$T\,dS = d\varepsilon + p\,dV \tag{27}$$

к жидкой частице.

Второе обстоятельство — особенная простота выражения энтропии через давление и плотность в случае идеального газа:

$$S = C_V \ln p\rho^{-\gamma} + S_0, \tag{28}$$

где $\gamma > 1$ — показатель адиабаты, равный отношению удельной теплоемкости при постоянном давлении (C_p) к постоянном объеме (C_V), S_0 — несущественная константа.

Из (26) с учетом (28) имеем

$$\frac{\partial(p\rho^{-\gamma})}{\partial t} = 0,$$

что эквивалентно выражению

$$p\rho^{-\gamma} = \varphi(m), \tag{29}$$

означающему независимость от времени энтропии любой частицы газа. Функция $\varphi(m)$ описывает распределение энтропии по массе газа, определяемое по заданным в момент $t = 0$ функциям $p_0(m)$, $\rho_0(m)$.

Используя интеграл (29) вместо дифференциального уравнения (20), сведем (18), (19) к дифференциальному уравнению второго порядка относительно плотности:

$$\frac{\partial^2\rho}{\partial t^2} = \frac{\partial}{\partial m}\left(a_0\rho^{\gamma+1}\frac{\partial\rho}{\partial m}\right), \tag{30}$$

где $a_0 = \gamma\varphi_0$, а постоянная φ_0 — энтропия, предполагаемая не зависящей также и от массовой координаты (случай изэнтропического течения).

Гиперболичность уравнения (30) и тем самым уравнений газовой динамики легко установить не прибегая к вычислению характеристик, а получив его линейный аналог. Для этого рассмотрим малые возмущения газодинамических величин в окрестности постоянного решения

$\rho(m,t) \equiv \rho_0$. Представляя возмущенное решение в виде $\rho(m,t) = \rho_0 + \bar\rho$ и предполагая малыми как сами возмущения, так и их производные, из (30) получаем уравнение для $\bar\rho$ (черточку опускаем):

$$\frac{\partial^2\rho}{\partial t^2} = c_0^2\frac{\partial^2\rho}{\partial m^2}. \tag{31}$$

Линейное уравнение (31) полностью аналогично уравнению колебаний струны из § 2 гл. III, имеющему, как известно, гиперболический тип. Оно описывает распространение малых (звуковых) возмущений в газе (*уравнение акустики*) со скоростью звука $c_0 = \sqrt{\gamma p_0/\rho_0}$ и, в силу линейности, для него нетрудно найти общее решение.

Еще одно упрощение уравнений (18), (19), (29) получается в предположении о том, что течение имеет характер простой волны, т. е. любые газодинамические величины являются функциями какой-то одной выбранной величины, например плотности. Из (18), (19), (29) и с учетом того, что $v = v(\rho)$, получаем

$$-\frac{1}{\rho^2}\frac{\partial\rho}{\partial t} = v_\rho\frac{\partial\rho}{\partial m}, \qquad v_\rho\frac{\partial\rho}{\partial t} = -a_0\rho^{\gamma-1}\frac{\partial\rho}{\partial m},$$

где v_ρ — производная скорости по плотности. Исключая из последних уравнений величину v_ρ, приходим к *уравнению Хопфа*

$$\frac{\partial\rho}{\partial t} + \sqrt{a_0}\,\rho^{\frac{\gamma+1}{2}}\frac{\partial\rho}{\partial m} = 0. \tag{32}$$

Уравнение (32) первого порядка, но оно содержит типичную газодинамическую нелинейность, и поэтому служит хорошей моделью для изучения нелинейных эффектов, характерных для течений сжимаемого газа. Самый яркий из них — «градиентная катастрофа», заключающаяся в появлении в волнах сжатия бесконечных градиентов газодинамических величин, несмотря на то, что в начальный момент времени все функции являются гладкими.

Поясним это понятие следующими простыми рассуждениями. Уравнение Хопфа может быть записано в характеристическом виде

$$\left(\frac{d\rho}{dt}\right)_s = 0. \tag{33}$$

Здесь индекс s означает, что полная производная по времени берется вдоль характеристики — линии в координатах m, t, на которой значение решения (плотности) остается постоянным во все моменты. Раскрывая (33) в виде

$$\frac{\partial\rho}{\partial t} + \frac{dm_s(t)}{dt}\frac{\partial\rho}{\partial m} = 0, \tag{34}$$

где $m_s(t)$ — значение координаты m для данной характеристики в разные моменты времени, и сравнивая (32) и (34), находим уравнение характеристики

$$m_s(t) = \sqrt{a_0}\,\rho^{\frac{\gamma+1}{2}}t + m_s(0).$$

Из этого выражения видно, что состояния с бо́льшим значением плотности распространяются по массе газа с большей скоростью, чем состояния с меньшей плотностью, и в какой-то момент времени «догоняют» последние. В решении образуется неоднозначность, его градиенты в точке «слияния» состояний с разными плотностями неограниченно возрастают. Схематически этот процесс изображен на рис. 36, где показана эволюция со временем начального профиля плотности треугольной формы: вершина треугольника через некоторое время оказывается в точке с той же координатой m_k, что и его передний фронт.

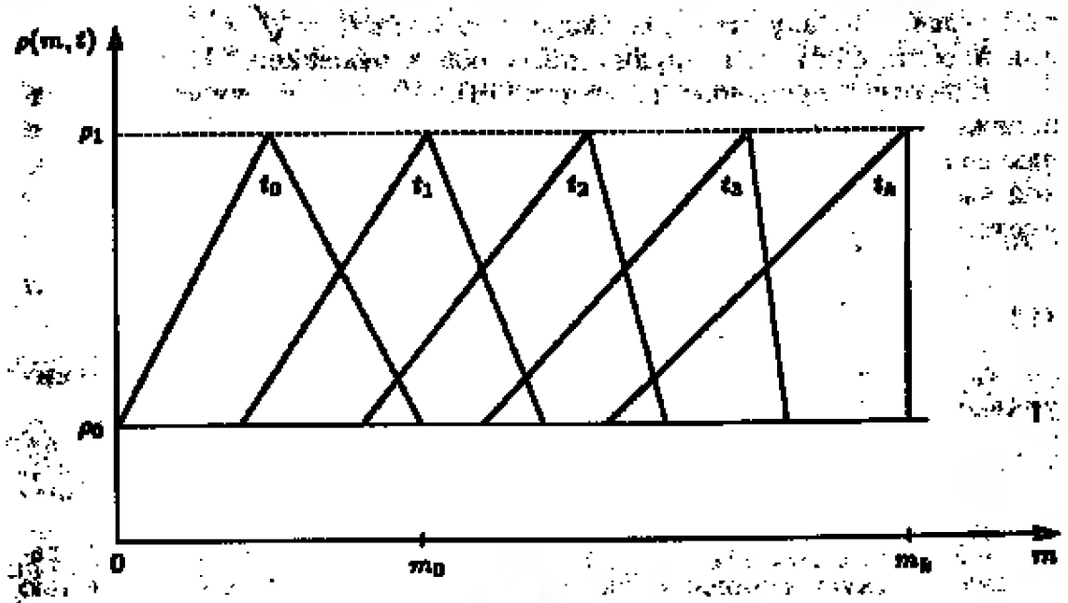


Рис. 36

«Градиентная катастрофа» — нелинейный эффект (подробнее см. гл. V). Он не возникает в линейном уравнении (31) и линейном уравнении переноса (3) § 1, получающемся из (32) при рассмотрении малых возмущений в окрестности постоянного решения. Существование этого эффекта приводит к необходимости рассматривать разрывные решения уравнений газовой динамики (заметим, что при прохождении разрыва через жидкую частицу ее энтропия изменяется). В этом состоит важное отличие нелинейных гиперболических уравнений от параболических (модели в § 1, 2).

УПРАЖНЕНИЯ

1. Получите уравнение (21) из (20), используя уравнение (19).
2. Используя закон сохранения энергии применительно к внутренней энергии жидкой частицы, выведите уравнение (22).
3. Убедитесь в эквивалентности уравнений (26) и (20), используя (27) и (28).
4. Получите уравнение (31) не из (30), а из (18), (19), (29), применяя те же рассуждения.
5. Используя характеристическую форму уравнения Хопфа, найдите момент t_k и точку m_k «градиентной катастрофы» для ситуации на рис. 36.

Библиография к главе II: [30, 43, 44, 55, 69, 74].

ГЛАВА III

МОДЕЛИ ИЗ ВАРИАЦИОННЫХ ПРИНЦИПОВ. ИЕРАРХИИ МОДЕЛЕЙ

§ 1. Уравнения движения, вариационные принципы и законы сохранения в механике

Приведем элементарные сведения из классической механики, иллюстрирующие связь между уравнениями динамики Ньютона и Лагранжа, вариационным принципом Гамильтона, законами сохранения для механических систем и свойствами пространства-времени.

1. Уравнения движения механической системы в форме Ньютона. Динамика системы, состоящей из N материальных точек, испытывающих действие каких-либо сил, описывается уравнениями

$$m_i \frac{d^2\vec{r}_i}{dt^2} = \vec{F}_i, \qquad i = 1, \ldots, N, \tag{1}$$

где m_i — масса материальной точки, $t > 0$ — время, $\vec{r}_i$ — ее радиус-вектор, $\vec{F}_i = \vec{F}_i(t, \vec{r}, d\vec{r}/dt)$ — результирующая всех действующих на нее сил. Через $\vec{r}$ обозначено множество координат всех точек системы. Величины $\vec{F}_i$ считаются заданными и могут зависеть как от времени, так и от пространственных координат и скоростей (причем не только выделенной i-й точки, но и всех рассматриваемых точек). Система (1) — математическое выражение второго закона Ньютона, примененного к совокупности материальных точек. Записанная в покоординатной форме она, очевидно, состоит из $3N$ уравнений второго порядка относительно $3N$ неизвестных координат $x_i(t)$, $y_i(t)$, $z_i(t)$, $i = 1, \ldots, N$. Если известны начальные координаты $x_i(0)$, $y_i(0)$, $z_i(0)$, $i = 1, \ldots, N$, точек и их скорости dx_i/dt, dy_i/dt, dz_i/dt в момент $t = 0$, то система (1) дает возможность найти координаты и скорости всех точек при любых $t > 0$, т. е. решить основную задачу классической механики.

Уравнения динамики в форме (1) справедливы для *инерциальной*, или *галилеевой системы отсчета*. В ней свободная материальная точка, т. е. точка, не испытывающая каких-либо воздействий, движется равномерно и прямолинейно. Другими словами, в этой системе координат выполняется первый закон Ньютона. Примером может служить система отсчета, жестко связанная с поездом, движущимся с постоянной скоростью по прямой и горизонтально ровной железнодорожной колее. Шарик, пущенный по гладкому полу вагона такого поезда, будет

двигаться относительно него равномерно и прямолинейно (это свойство сразу же утрачивается при торможениях и ускорениях поезда, на поворотах, подъемах и спусках — система отсчета становится неинерциальной). Конечно же, инерциальная система отсчета — идеализированное понятие, оправданное, однако, при рассмотрении многих важных механических явлений. Очевидное отсутствие инерциальности системы отсчета, жестко связанной с Землей (хотя бы из-за ее вращения), не мешает с хорошей точностью описывать ряд «земных» движений таких, как движение поезда по ее поверхности (для полета баллистической ракеты подобное описание может быть уже неточным).

Любая система отсчета, покоящаяся или прямолинейно движущаяся с постоянной скоростью по отношению к некоторой инерциальной системе, также является инерциальной. Множество таких систем, «порождаемых» исходной системой x, y, z, t, определяется следующими преобразованиями координат и времени:

$$
\begin{aligned}
&x^* = x + a, \quad y^* = y + b, \quad z^* = z + c, \quad t^* = t,\\
&x^* = x, \quad y^* = y, \quad z^* = z, \quad t^* = t + a,\\
&x^* = x\cos\alpha + y\sin\alpha, \quad y^* = -x\sin\alpha + y\cos\alpha, \quad z^* = z, \quad t^* = t,\\
&x^* = x - v_x t, \quad y^* = y - v_y t, \quad z^* = z - v_z t, \quad t^* = t,
\end{aligned}
\tag{2}
$$

а также любыми комбинациями этих преобразований. Здесь a, b, c, v_x, v_y, v_z — произвольные постоянные так же, как и α — угол поворота системы отсчета относительно одной из осей координат (в данном случае оси z).

Классическая механика исходит из принципа относительности Галилея — равноправия всех инерциальных систем (2), т. е. из того, что законы механики одинаковы в любой из них. Этот принцип отражает, в частности, однородность пространства и времени (первое и второе преобразования (2)) и изотропность пространства — отсутствие в нем выделенных направлений (третье преобразование). Последнее из преобразований (2) — переход к системе отсчета, движущейся с постоянной скоростью относительно исходной, — называется *преобразованием Галилея*.

Заметим, что если по каким-то причинам приходится рассматривать движение в заведомо неинерциальной системе отсчета, то второй закон Ньютона в форме (1) не выполняется. Так, шарик, пущенный по гладкой горизонтальной поверхности вращающегося диска от центра к его краю, в системе отсчета, жестко связанной с диском, движется уже не прямолинейно, а по кривой траектории, как бы испытывая действие некоторой внешней силы. Для неинерциальных систем уравнения (1) и другие законы механики становятся справедливыми при их относительно несложной модификации. Ее суть заключается в добавлении к действующим на систему силам «фиктивных» внешних *сил инерции*, величина которых определяется из характера движения неинерциальной системы отсчета по отношению к выбранной инерциальной

системе. Обычно же исходная система отсчета в механике считается инерциальной, а ее неинерциальность специально оговаривается.

Неизменность законов механики по отношению к преобразованиям (2) может выражаться различным образом. Если для уравнений, соответствующих этим законам, при переходе к новой системе отсчета

1) не меняется их структура,

2) не меняется вид функций от координат, скоростей и ускорений, фигурирующих в уравнениях (ими являются силы, энергия, количество движения и другие механические величины),

то уравнения *инвариантны* по отношению к заданному преобразованию. Примером служат уравнения

$$
\begin{aligned}
m_1 \frac{d^2 r_1}{dt^2} &= -k(l + r_1 - r_2),\\
m_2 \frac{d^2 r_2}{dt^2} &= -k(r_2 - r_1 - l),
\end{aligned}
\tag{3}
$$

неизменные при любом преобразовании (2) (см. также упр. 1).

Уравнение, описывающее колебания колец Сатурна, из п. 3 § 2 гл. I

$$
M_2 \frac{d^2 r}{dt^2} = \gamma M_1 M_2 \frac{r}{(r^2 + R_0^2)^{3/2}},
$$

инвариантно при «сдвиге» времени $t^* = t + a$, но изменяет свою форму при «сдвиге» координаты $r^* = r + a$:

$$
M_2 \frac{d^2 r^*}{dt^2} = \gamma M_1 M_2 \frac{r^* - a}{((r^* - a)^2 + R_0^2)^{3/2}},
$$

в том смысле, что выражение, стоящее в его правой части (т. е. сила), как функция координаты записывается иначе, чем в исходном уравнении. Форма записи уравнений механики, теряющая свойство 1), но сохраняющая свойство 2) при преобразованиях (2), называется *ковариантной*. Из покоординатной записи уравнений (1) непосредственно следует их ковариантность относительно преобразований (2) (однако при более сложных преобразованиях исходной системы отсчета, например, при переходе к цилиндрическим и сферическим координатам, ковариантность уравнений движения в форме Ньютона утрачивается).

2. Уравнения движения в форме Лагранжа. Уравнения Лагранжа — удобная запись второго закона Ньютона, ковариантная по отношению к гораздо более широкому классу преобразований систем отсчета, нежели (2) (ковариантность понимается в том же смысле, что и при преобразованиях инерциальных систем). Поясним, как получаются эти уравнения, на примере самой простой механической системы, состоящей из одной материальной точки, совершающей одномерное движение. В исходной инерциальной системе координат оно описывается уравнением

$$
m \frac{d^2 r}{dt^2} = F\left(r, \frac{dr}{dt}, t\right). \tag{4}
$$

Пусть задано преобразование координат $r(t) = r(q(t), t)$, где $q(t)$ — новая координата материальной точки. Для ее скорости получаем следующее выражение через функции $r(q)$ и $q(t)$:

$$v = \frac{dr}{dt} = \frac{\partial r}{\partial q}\dot{q} + \frac{\partial r}{\partial t}. \tag{5}$$

Здесь $\dot{q} = dq/dt$ — «скорость» точки в новых координатах (это в общем случае уже не реальная кинематическая скорость, а всего лишь характеристика темпа изменения координаты q точки со временем). «Настоящая» же скорость v в новой системе отсчета, в отличие от старой, становится согласно (5) функцией координаты q и скорости $\dot{q}$, т. е. $v = v(q, \dot{q}, t)$. Вычислим теперь ускорение точки ω в новых координатах, представив его предварительно в эквивалентном виде:

$$\omega = \frac{d^2r}{dt^2} = \frac{dv}{dt} = \frac{1}{\partial r/\partial q}\frac{dv}{dt}\frac{\partial r}{\partial q} = \frac{1}{\partial r/\partial q}\left[\frac{d}{dt}\left(v\frac{\partial r}{\partial q}\right) - v\frac{d}{dt}\frac{\partial r}{\partial q}\right].$$

Продифференцировав (5) по $\dot{q}$, получим

$$\frac{\partial r}{\partial q} = \frac{\partial v}{\partial \dot{q}},$$

т. е. выражение для преобразования первого члена в квадратных скобках. Для второго члена, учитывая (5), имеем цепочку равенств

$$\frac{d}{dt}\frac{\partial r(q(t),t)}{\partial q} = \frac{\partial^2 r}{\partial q^2}\dot{q} + \frac{\partial^2 r}{\partial q \partial t} = \frac{\partial}{\partial q}\left[\frac{\partial r}{\partial q}\dot{q} + \frac{\partial r}{\partial t}\right] = \frac{\partial v}{\partial q}.$$

Используя два последних равенства, придем к окончательному выражению

$$\omega(q, \dot{q}, t) = \frac{1}{\partial r/\partial q}\left[\frac{d}{dt}\frac{\partial(v^2/2)}{\partial \dot{q}} - \frac{\partial(v^2/2)}{\partial q}\right],$$

подставив которое в (4), получим уравнение

$$\frac{d}{dt}\frac{\partial(mv^2/2)}{\partial \dot{q}} - \frac{\partial(mv^2/2)}{\partial q} = F(q, \dot{q}, t)\frac{\partial r}{\partial q}. \tag{6}$$

В (6) величина $F(q, \dot{q}, t)$ — сила, стоящая в правой части уравнения (4), но записанная в новых координатах, зависимость ее от аргументов q, $\dot{q}$, t в общем случае не такая, как от r, $\dot{r}$, t. Уравнение (6) — простейшее уравнение движения в форме Лагранжа. В его левой части под знаками дифференцирования стоит, очевидно, кинетическая энергия точки, а в правой — действующая на нее сила, умноженная на величину $\partial r/\partial q$, характеризующую связь между старыми и новыми координатами.

При трехмерном движении точки вместо координат x, y, z необходимо ввести три новые координаты q_1, q_2, q_3. Повторим рассуждения, предшествовавшие выводу (6), но возьмем в качестве v, ω, F поочередно v_x, v_y, v_z; ω_x, ω_y, ω_z; F_x, F_y, F_z — соответственно проекции скорости, ускорения и силы. Рассмотрим поочередно их как функции

аргументов $q_1, \dot{q}_1, t$; $q_2, \dot{q}_2, t$; $q_3, \dot{q}_3, t$ и подставим полученные выражения в записанные покоординатно уравнения Ньютона $m\,dv_x/dt = F_x$, $m\,dv_y/dt = F_y$, $m\,dv_z/dt = F_z$.

В результате придем уже к трем уравнениям, аналогичным (6). Например, первое из них имеет вид

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_1} - \frac{\partial T}{\partial q_1} = F_x\frac{\partial x}{\partial q_1} + F_y\frac{\partial y}{\partial q_1} + F_z\frac{\partial z}{\partial q_1}, \tag{7}$$

где $T = m(v_x^2 + v_y^2 + v_z^2)/2$ — кинетическая энергия точки.

Из этих примеров ясна процедура для общего случая движения N точек, положение каждой из которых определяется тремя координатами $x_i(t)$, $y_i(t)$, $z_i(t)$, $i = 1, \ldots, N$ (выкладки вполне схожи с теми, что использовались при получении (6), (7), и не приводятся ввиду их громоздкости). Переход к новым координатам задается произвольным преобразованием

$$\begin{aligned} x_i &= x_i(q_1, \ldots, q_n; t), \\ y_i &= y_i(q_1, \ldots, q_n; t), \\ z_i &= z_i(q_1, \ldots, q_n; t), \end{aligned} \tag{8}$$

где $i = 1, \ldots, N$, $n = 3N$. В (8) величины q_j, $1 \leqslant j \leqslant 3N$, — так называемые *обобщенные координаты*. Их общее число, как и в старой системе отсчета, естественно, равно $3N$. При переходе к координатам q_j любая старая координата может зависеть, вообще говоря, от всех новых (как, например, в случае перехода от декартовых координат x, y, z к сферическим координатам ρ, φ, ψ). Поэтому в правых частях формул (8) предусмотрена зависимость от всех величин q_j.

После подсчета «проекций» (точнее, компонент) всех векторов $\vec{v}_i$, $\vec{\omega}_i$, $\vec{F}_i$, $i = 1, \ldots, N$, на все новые «координатные оси» q_j, подстановки полученных выражений в покоординатные уравнения (1) и их суммирования по индексу i при фиксированном значении j приходим к общим *уравнениям Лагранжа*

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} = \Phi_j, \quad j = 1, \ldots, n. \tag{9}$$

В (9), как и при $N = 1$, величина T — кинетическая энергия системы

$$T = \sum_{i=1}^{N}\frac{m_i v_i^2}{2} = \sum_{i=1}^{N}\frac{m_i \vec{v}_i \cdot \vec{v}_i}{2}, \quad i = 1, \ldots, N,$$

записанная в координатах q_j. Величины

$$\Phi_j = \sum_{i=1}^{N}\left(F_{ix}\frac{\partial x_i}{\partial q_j} + F_{iy}\frac{\partial y_i}{\partial q_j} + F_{iz}\frac{\partial z_i}{\partial q_j}\right) = \sum_{i=1}^{N}\vec{F}_i(q, \dot{q}, t)\frac{\partial \vec{r}_i}{\partial q_j} \tag{10}$$

можно понимать как выраженные в координатах q_j «проекции» сил $\vec{F}_i$ на оси q_j — *обобщенные силы* (функции $|\partial \vec{r}_i/\partial q_j|$ — *коэффициенты*

Ламе; через $\vec{A}\cdot\vec{B}$ обозначено скалярное произведение векторов $\vec{A}$ и $\vec{B}$). В указанной трактовке уравнения (9) представляют собой запись второго закона Ньютона (1) в «проекциях» на оси q_j.

Они заметно упрощаются в случае, когда все силы потенциальны, т. е. существует функция $\Pi(x_1,\ldots,x_N;y_1,\ldots,y_N;z_1,\ldots,z_N;t)$ такая, что

$$F_{ix}=-\frac{\partial\Pi}{\partial x_i},\quad F_{iy}=-\frac{\partial\Pi}{\partial y_i},\quad F_{iz}=-\frac{\partial\Pi}{\partial z_i},\qquad i=1,\ldots,N.$$

Подставим эти выражения в (10) и получим

$$\Phi_j=-\sum_{i=1}^{N}\left(\frac{\partial\Pi}{\partial x_i}\frac{\partial x_i}{\partial q_j}+\frac{\partial\Pi}{\partial y_i}\frac{\partial y_i}{\partial q_j}+\frac{\partial\Pi}{\partial z_i}\frac{\partial z_i}{\partial q_j}\right).$$

Заменяя в последнем равенстве в функции Π старые аргументы на новые, приходим к выводу, что его правая часть представляет собой частную производную некоторой функции $V(q,t)$, называемой *потенциалом*, по аргументу q_j, т. е.

$$\Phi_j=-\frac{\partial V}{\partial q_j}.$$

Другими словами, если исходные силы потенциальны, то и обобщенные силы потенциальны, и уравнения (9) приобретают вид

$$\frac{d}{dt}\frac{\partial L}{\partial\dot q_j}-\frac{\partial L}{\partial q_j}=0,\qquad j=1,\ldots,n, \tag{11}$$

$$L=T-V. \tag{12}$$

При получении (11), (12) учтено, что функция $V(q,t)$ не зависит от $\dot q$, и поэтому $\partial L/\partial q_j=\partial T/\partial\dot q_j$. В уравнениях Лагранжа (11) величина L — разность кинетической и потенциальной энергий системы, выраженных в новых координатах. Ее называют *функцией Лагранжа*, *лагранжианом* или *кинетическим потенциалом системы*.

Уравнения (9), (11), полученные из уравнений Ньютона (1), ковариантны, как следует из их вида, относительно произвольного преобразования (8) исходной системы отсчета, а входящие в них функции имеют ясный механический смысл. Для их написания применительно к конкретной системе необходимо выполнить следующие действия: выбрать независимые координаты $q_1,\ldots,q_n$; найти обобщенные силы (10) как функции новых координат (если силы потенциальны, то этого делать не нужно); выразить в новых координатах кинетическую энергию T и потенциальную энергию Π и найти лагранжиан (12) (в случае потенциальных сил); подставить полученные выражения в (9), (11).

Эта стандартная процедура называется *лагранжевым формализмом*. После ее реализации получается система из $3N$ дифференциальных уравнений второго порядка для координат $q_1,\ldots,q_n$, всегда разрешимая относительно их вторых производных:

$$\ddot q_j=G_j(q,\dot q,t),\qquad j=1,2,\ldots,n. \tag{13}$$

При известных начальных значениях $q_j(0)$, $\dot q_j(0)$, $j=1,\ldots,n$, уравнения (13) дают возможность описать движение системы в любой момент $t>0$, т. е. решить основную задачу механики.

Преимущества лагранжева подхода заключаются в его единообразии (ковариантность), в получении относительно простого математического описания движения очень сложных механических устройств (например, робототехнических) в произвольных системах отсчета (в том числе, как видно из (8), и в неинерциальных системах без введения дополнительных сил инерции). Заметим, что уравнения (1) также могут быть выписаны в виде, ковариантном по отношению к преобразованиям (8). Однако они не разрешаются относительно старших производных и содержат гораздо больше (по сравнению с уравнением (9)) функций от новых координат, не имеющих к тому же простого механического смысла.

Еще одно преимущество описанного формализма проявляется при изучении систем с *механическими связями* (во всех предыдущих рассуждениях для упрощения рассматривались свободные системы). Связи накладывают на движение точек системы определенные ограничения, вызванные присутствием материальных объектов, непосредственно в нее не входящих. Примерами служат связи из-за наличия жесткой поверхности, ограничивающей движение катящегося по ней шарика, или невесомого жесткого стержня, соединяющего две точечные массы, и т. п.

Связи задаются набором K соотношений вида $f_k(\vec r_i,\dot{\vec r},t)=0$, $0\leqslant k\leqslant K\leqslant 3N$, которым должно удовлетворять решение уравнений Ньютона (1). Сами же уравнения (1) в этом случае модифицируются, в их правых частях следует (при некоторых предположениях о характере связей) подставить вместо $\vec F_i$ величину $\vec F_i+\vec R_i$. Сила $\vec R_i$ — вычисляемая из тех или иных соображений *реакция связей*. Поэтому при подходе Ньютона описание несвободных систем по сравнению со свободными может заметно усложниться.

При подходе Лагранжа реакции связей для очень широкого класса функций $f_k(\vec r_i,\dot{\vec r},t)$ учитываются автоматически, в результате чего обобщенные силы (10) не содержат в себе величин $\vec R_i$ (сама процедура получения уравнений Лагранжа полностью аналогична случаю свободных систем). Более того, число уравнений Лагранжа равно количеству так называемых степеней свободы $l=3N-K$ и может быть значительно меньше, чем для системы без связей. Число новых независимых координат q_j, естественно, также равно $3N-K$.

Итак, уравнения Лагранжа выводятся из уравнений Ньютона, и наоборот (упр. 2). Таким образом, лагранжев подход, как и подход Ньютона, может быть положен в основу механики.

3. Вариационный принцип Гамильтона. В качестве фундамента механики могут рассматриваться не только дифференциальные уравнения (1), (9) или (11), связывающие механические параметры в данный момент времени t, но также некоторые общие свойства, характеризующие движение механической системы в целом, на любом произвольном отрезке времени от t_0 до t_1. Убедимся в этом, анализи-

руя величину

$$Q = \int_{t_0}^{t_1} L(q(t), \dot{q}(t), t)\, dt, \tag{14}$$

называемую *действием по Гамильтону на отрезке* $[t_0, t_1]$. Очевидно, что (14) является функционалом, зависящим от того, как движется система в моменты $t_0 \leqslant t \leqslant t_1$.

Возьмем в $(n+1)$-мерном пространстве q, t две точки $M_0(q(t_0), t_0)$ и $M_1(q(t_1), t_1)$, зафиксировав тем самым моменты t_0, t_1 и положение системы в эти моменты («скорости» $\dot{q}$ в моменты t_0, t_1 не фиксируются). Из точки M_0 в точку M_1 система может попасть, двигаясь в пространстве q, t, вообще говоря, по любым кинематически возможным траекториям («путям»), т. е. по путям, допускаемым существующими связями (см. рис. 37 для случая пространства q_1, q_2, t). Пусть среди этих путей имеется так называемый *прямой путь* (сплошная линия). На нем функции $q_j(t)$, $j = 1, \ldots, n$, в любой момент времени подчиняются уравнениям Лагранжа (11). Остальные пути называются *окольными* (штриховые линии).

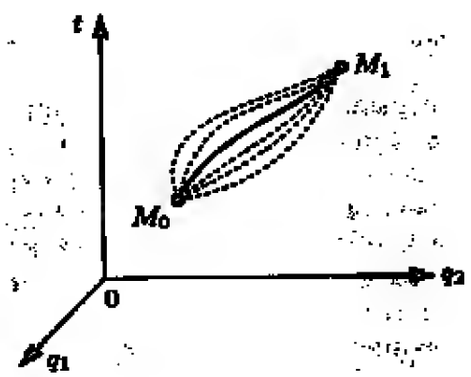


Рис. 37

Принцип Гамильтона формулируется так: действие Q имеет на прямом пути экстремальное по сравнению с окольными путями значение.

Охарактеризуем все возможные пути однопараметрическим семейством функций

$$q_j = q_j(t, \alpha), \qquad t_0 \leqslant t \leqslant t_1, \quad |\alpha| \leqslant \beta < \infty, \quad j = 1, \ldots, n,$$

где значению $\alpha = 0$ отвечает прямой путь, а значениям $\alpha \neq 0$ — окольные пути. Тогда действие (14), очевидно, является функцией параметра α:

$$Q(\alpha) = \int_{t_0}^{t_1} L[q_j(t, \alpha), \dot{q}_j(t, \alpha), t]\, dt.$$

Вариация Q при варьировании параметра α есть

$$\delta Q = \frac{\partial q}{\partial \alpha}\, d\alpha = \int_{t_0}^{t_1} \delta L\, dt = \int_{t_0}^{t_1} \sum_{j=1}^{n} \left(\frac{\partial L}{\partial q_j}\, \delta q_j + \frac{\partial L}{\partial \dot{q}_j}\, \delta \dot{q}_j \right) dt, \tag{15}$$

т. е. равна сумме приращений, вызванных вариацией координат $\delta q_j(t, \alpha)$ и скоростей $\delta \dot{q}_j(t, \alpha)$.

Проинтегрировав по частям второй член в правой части (15), получаем

$$\delta Q = \sum_{j=1}^{n} \frac{\partial L}{\partial \dot{q}_j}\, \delta(\dot{q}_j) \Bigg|_{t_0}^{t_1} - \int_{t_0}^{t_1} \sum_{j=1}^{n} \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} \right) \delta q_j\, dt,$$

или, используя перестановочность операций варьирования по α и дифференцирования по t:

$$\delta(\dot{q}_j) = \delta \left[\frac{dq_j(t, \alpha)}{dt} \right] = \frac{\partial}{\partial \alpha} \frac{d}{dt}\, q_j(t, \alpha)\, \delta\alpha = \frac{d}{dt} \left[\frac{\partial}{\partial \alpha}\, q_j(t, \alpha)\, \delta\alpha \right] = \frac{d}{dt}\, \delta q_j,$$

приходим к равенству

$$\delta Q = \sum_{j=1}^{n} \frac{\partial L}{\partial \dot{q}_j}\, \delta q_j \Bigg|_{t_0}^{t_1} - \int_{t_0}^{t_1} \sum_{j=1}^{n} \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} \right) \delta q_j\, dt. \tag{16}$$

По построению вариации $\delta q_j(\alpha)$ равны нулю в моменты t_0, t_1, т. е. равен нулю первый член в правой части (16). Для прямого пути по определению справедливы уравнения Лагранжа (11), поэтому равен нулю также и второй член.

Таким образом, на прямом пути $\delta Q = 0$. В этом состоит математическое выражение принципа Гамильтона. Его еще называют *принципом наименьшего действия*, поскольку действие вдоль прямого пути, как можно показать при некоторых дополнительных предположениях, имеет наименьшее в сравнении с окольными путями значение. Для более общих механических систем (не подчиняющихся уравнениям (11)) схожее утверждение называется *принципом Гамильтона—Остроградского*.

Напомним, что в терминах вариационного исчисления уравнения (11) называются *дифференциальными уравнениями Эйлера* для вариационной задачи

$$\delta Q = \delta \int_{t_0}^{t_1} L(q, \dot{q}, t)\, dt = 0,$$

прямой путь — *экстремалью*, а отвечающее ему число $Q(\alpha = 0)$ — *стационарным значением функционала* Q.

Нетрудно убедиться в справедливости утверждения, обратного принципу Гамильтона: если для некоторого пути выполняется свойство $\delta Q = 0$, то этот путь является прямым (упр. 3). Поэтому все три рассмотренных выше подхода равноправны и могут быть положены в основу математического описания механических систем.

4. Законы сохранения и свойства пространства-времени. Анализ интегральных механических характеристик не сводится к формулировке принципа Гамильтона и различных его обобщений. Он позволяет установить и другие фундаментальные свойства механических систем. Продемонстрируем некоторые из этих свойств в случае

движений в потенциальных полях, когда как внешние, так и внутренние силы, действующие на точки системы, потенциальны. Для этого понадобится вычислять вариацию действия Q на более общем, нежели совокупность прямых и окольных путей из п. 3, семействе кривых в $(n+1)$-мерном пространстве $q_1, \ldots, q_n, t$.

Рассмотрим это однопараметрическое семейство

$$q_1 = q_1(t, \alpha), \quad \ldots, \quad q_n = q_n(t, \alpha), \tag{17}$$

где α — параметр, однозначно определяющий кривую, в частности ее «нижний» и «верхний» концы в пространстве q, t (рис. 38). На пучок (17) не накладывается каких-либо ограничений, поэтому кривые могут «начинаться» и «кончаться» в разные моменты времени $t_0(\alpha)$ и $t_1(\alpha)$, пересекаться между собой, разные кривые могут совпадать в начальные и конечные моменты и т. д. Вариация действия

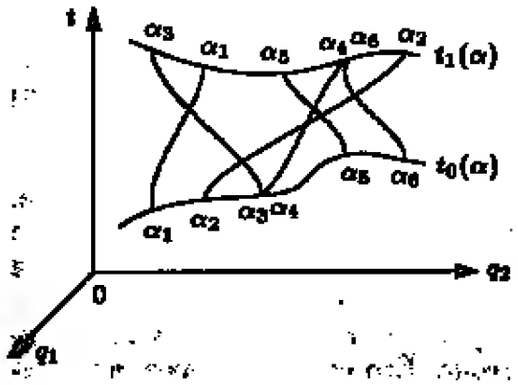


Рис. 38

$$\delta Q = \delta \int\limits_{t_0(\alpha)}^{t_1(\alpha)} L(q, \dot q, t)\, dt$$

на пучке (17) находится с помощью рассуждений, вполне аналогичных тем, что использовались при получении (16), однако соответствующие выкладки гораздо сложнее, так как начальное и конечное положения кривых семейства в пространстве q, t зависят от параметра α. Поэтому приведем только конечный результат: общая формула для δQ в случае однопараметрического пучка (17) имеет вид

$$\delta Q = \left(\sum_{j=1}^{n} p_j\, \delta q_j - H\, \delta t\right)\Bigg|_{t_0(\alpha)}^{t_1(\alpha)} - \int\limits_{t_0(\alpha)}^{t_1(\alpha)} \sum_{j=1}^{n} \left(\frac{d}{dt}\frac{\partial L}{\partial \dot q_j} - \frac{\partial L}{\partial q_j}\right) \delta q_j\, dt. \tag{18}$$

Смысл функций p_j и H будет разъяснен ниже.

Очевидно, что если начала и концы всех кривых пучка совпадают, то $\delta q_j = \delta t = 0$, и из (18) сразу следует (16), или принцип Гамильтона.

Функции p_j и функция H называются соответственно *обобщенными импульсами* и *гамильтонианом системы*. Величины p_j вводятся как частные производные лагранжиана по скоростям $\dot q_j$:

$$p_j = \frac{\partial L}{\partial \dot q_j} = p_j(q, \dot q, t), \qquad j = 1, \ldots, n, \tag{19}$$

и в простейшем случае движения точки в декартовых координатах в стационарном потенциальном поле (функция Π из п. 2 не зависит явно от времени) совпадают с проекциями количества движения на оси x, y, z. В силу не обсуждаемых здесь общих свойств лагранжиана

соотношения (19) могут быть разрешены относительно обобщенных скоростей:

$$\dot q_j = h_j(q, p, t),$$

т. е. существует взаимно однозначный переход от лагранжевых переменных q, $\dot q$, t к так называемым *гамильтоновым переменным* q, p, t. Функция H определяется равенством

$$H(q, p, t) = \sum_{j=1}^{n} p_j \dot q_j - L, \tag{20}$$

где лагранжиан L и обобщенные скорости выписаны в переменных q, p, t. Гамильтониан играет важную роль при изучении потенциальных движений, в частности потому, что из уравнений Лагранжа (11) получаются уравнения движения в форме Гамильтона (канонические уравнения)

$$\dot q_j = \frac{\partial H}{\partial p_j}, \quad \dot p_j = -\frac{\partial H}{\partial q_j}, \qquad j = 1, \ldots, n, \tag{21}$$

и наоборот (упр. 4). Функция H имеет ясный механический смысл: если преобразования исходной системы (8) стационарны (явно не зависят от t), то в любой момент времени гамильтониан численно равен полной энергии системы, т. е.

$$H = T + \Pi = E.$$

После этих пояснений установим связь законов сохранения в механике со свойствами пространства и времени. Рассмотрим однопараметрическое преобразование системы отсчета q, t

$$q_j^* = \varphi_j(q, t, \alpha), \qquad j = 1, \ldots, n, \qquad t^* = \psi(q, t, \alpha) \tag{22}$$

такое, что оно тождественно при $\alpha = 0$ и для него существует обратное преобразование.

Пусть лагранжиан заданной механической системы, движущейся в потенциальном поле, инвариантен относительно преобразований (22). Это значит, что новый лагранжиан $L^*(q^*, \dot q^*, t^*)$ не зависит от α и как функция q^*, $\dot q^*$, t^* имеет тот же вид, что и исходный лагранжиан в переменных q, $\dot q$, t. Тогда существует функция от p, q, t

$$\phi(q, p, t) = \sum_{j=1}^{n} p_j \left(\frac{\partial \varphi_j}{\partial \alpha}\right)_{\alpha=0} - H\left(\frac{\partial \psi_j}{\partial \alpha}\right)_{\alpha=0}, \tag{23}$$

не меняющая на прямых путях своего значения с течением времени (первый интеграл движения). В (23) p_j, H — соответственно обобщенные импульсы (19) и гамильтониан (20), записанные в переменных p, q, t.

Схема доказательства сформулированного утверждения (теорема Нётер) выглядит так. В пространстве q, t выбирается кривая $q(t)$, на которой $\delta Q = 0$, т. е. отвечающая части некоторого прямого пути на

отрезке $[t_0, t_1]$. В соответствии с (22) эта кривая «порождает» в пространстве q^*, t^* семейство кривых $q^*(t^*, \alpha)$. На них, в силу инвариантности лагранжиана, вариация действия равна нулю при всех значениях α, или, по формуле (18), примененной для пространства q^*, t^*,

$$\delta Q = \left(\sum_{j=1}^{n} p_j^* \, \delta q_j^* - H^* \, \delta t^*\right)\Bigg|_{t_0^*(\alpha)}^{t_1^*(\alpha)} - \int\limits_{t_0^*(\alpha)}^{t_1^*(\alpha)} \left(\sum_{j=1}^{n} \frac{d}{dt^*}\frac{\partial L^*}{\partial \dot q^*} - \frac{\partial L^*}{\partial q_j^*}\right) \delta q_j^* \, dt = 0. \qquad (24)$$

Если в (24) положить $\alpha = 0$, т. е. взять тождественное преобразование, то подынтегральное выражение обращается в нуль, так как в координатах q, t кривая $q(t)$ — прямой путь, и на нем удовлетворяются уравнения Лагранжа (11) для $L = L^*$. Значит, необходимо, чтобы выполнялось

$$\left[\left(\sum_{j=1}^{n}(p_j^* \, \delta q_j^* - H^* \, \delta t^*)\right)\Bigg|_{t_0^*(\alpha)}^{t_1^*(\alpha)}\right]_{\alpha=0} = 0.$$

Отсюда с учетом свойств преобразования (22) нетрудно вычислить δq_j^* и δt^* и, устремив α к нулю, получить формулу (23). Поскольку прямой путь и точки t_0, t_1 были выбраны произвольно, то утверждение теоремы справедливо для любого прямого пути (для всех истинных движений) системы.

Из теоремы Нётер следует закон сохранения полной механической энергии для систем, лагранжиан (так же как и гамильтониан) которых явно не зависит от времени, или *консервативных систем*. Действительно, взяв в качестве (22) сдвиг по времени

$$q_j^* = q_j, \qquad j = 1, \ldots, n, \qquad t^* = t + \alpha,$$

легко убедиться в инвариантности лагранжиана и из (23) получить

$$-\phi = H = T + \Pi = \text{const}.$$

Аналогичным образом устанавливаются закон сохранения количества движения для замкнутой (не испытывающей действия внешних сил) системы (упр. 5) и ряд других законов механики. При рассмотрении движений не на одном прямом пути, а на некоторых выделенных множествах прямых путей получаются более общие первые интегралы механических систем — интегральные инварианты. Некоторые из них могут быть положены, как и принцип Гамильтона, в основу механики.

Глубокая связь уравнений движения, законов сохранения, вариационных принципов и свойств симметрии позволяет эквивалентным образом использовать разнообразные подходы для построения математических моделей механических систем. Заметим, что инвариантные

свойства объектов эффективно применяются не только для построения, но и для анализа (см., например, гл. V) моделей многих явлений.

УПРАЖНЕНИЯ

1. Покажите, что уравнения (3) описывают движение шариков с массами m_1, m_2, соединенных невесомой пружиной с жесткостью k (l — длина ненагруженной пружины, $r_1(t) \leqslant r_2(t)$). Проверьте инвариантность уравнений при переходе от одной инерциальной системы к другой.

2. Взяв в качестве (8) тождественное преобразование исходной системы отсчета, убедитесь в том, что из уравнений (9) получается неинвариантная форма записи уравнений (1).

3. Проверьте, используя равенство (16), что для пути, вдоль которого $\delta Q = 0$, удовлетворяются уравнения Лагранжа (11).

4. Установите эквивалентность уравнений Гамильтона (21) и Лагранжа (11) на примере движения одной материальной точки в стационарном потенциальном поле, рассматривая его в декартовых координатах x, y, z.

5. В отсутствие внешних сил потенциальная энергия системы, а значит, и функция Лагранжа не меняются при сдвиге начала координат. Взяв в качестве (22) преобразование декартовых координат $x_i^* = x_i + \alpha$, $y^* = y_i$, $z^* = z_i$ $(i = 1, \ldots, N)$, $t^* = t$, убедитесь в том, что (23) имеет вид $\phi = \sum_{i=1}^{N} m_i \dot x_i = \text{const}$, т. е. что для такой системы справедлив закон сохранения количества движения в проекции на ось x.

§ 2. Модели некоторых механических систем

Уравнения Лагранжа и принцип Гамильтона применим для описания различных типов движения маятника и малых колебаний струны, а также колебаний электрического тока в контуре, для чего используем электромеханическую аналогию. Обсудим некоторые свойства изучаемых процессов.

1. Маятник на свободной подвеске. Система состоит из двух точечных масс, m_1 и m_2, соединенных невесомым жестким стержнем длины l (рис. 39). Движение происходит в поле силы тяжести и считается плоским, т. е. рассматривается в системе отсчета x, y, t. Точка с массой m_1 (подвеска) не закреплена, а может перемещаться вдоль оси x (ср. с п. 3 § 3 гл. I).

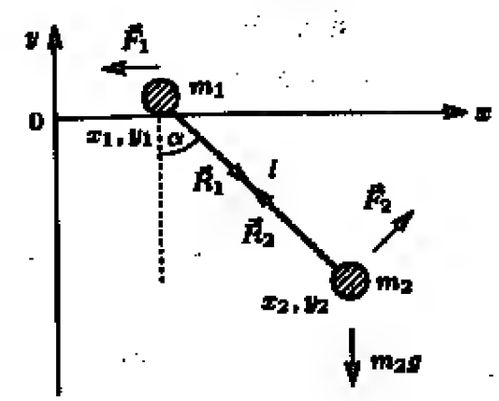


Рис. 39

Для описания плоского движения двух точек в исходной системе отсчета необходимо, вообще говоря, найти из уравнений (1) § 1 четыре функции времени, $x_1(t)$, $y_1(t)$, $x_2(t)$, $y_2(t)$, т. е. декартовы координаты первой и второй точек. Однако изучаемая система несвободна, поскольку содержит две механические связи (п. 2 § 1). Одна из них описывается уравнением $y_1 \equiv 0$ (подвеска не может совершать вертикальных перемещений), а вторая — уравнением $(x_1 - x_2)^2 + y_2^2 = l^2$ (расстояние между точками при любом t

равно длине стержня). Поэтому при переходе к уравнениям Лагранжа достаточно выбрать (по числу степеней свободы) лишь две новые независимые координаты. Возьмем в качестве обобщенных координат величины $q_1(t) = x_1(t)$ и $q_2(t) = \alpha(t)$, где α — угол между вертикалью и осью стержня. Такой выбор соответствует преобразованию (8) § 1 вида

$$x_1 = q_1, \quad x_2 = q_1 + l\sin q_2, \quad y_2 = -l\cos q_2.$$

Выразим сначала кинетическую энергию системы $T = T_1 + T_2$ в координатах q_1, q_2. Для подвески имеем

$$T_1 = \frac{m_1 v_1^2}{2} = \frac{m_1 v_{1x}^2}{2} = \frac{m_1 \dot{x}_1^2}{2}.$$

Для маятника получаем

$$T_2 = \frac{m_2 v_2^2}{2} = \frac{m_2}{2}(v_{2x}^2 + v_{2y}^2).$$

С помощью равенств $v_{2x} = \dot{x}_1 + l\dot{\alpha}\cos\alpha$, $v_{2y} = l\dot{\alpha}\sin\alpha$, первое из которых учитывает составное движение массы m_2 вдоль оси x как сумму движений вместе с подвеской и относительно нее, запишем величину T_2 как функцию x_1, α:

$$T_2 = \frac{m_2 \dot{x}_2^2}{2} + \frac{m_2}{2}(2l\dot{\alpha}\dot{x}_1\cos\alpha + l^2\dot{\alpha}^2).$$

Рассмотрим теперь силы, действующие на точки m_1, m_2. Сила тяжести и вертикальная проекция R_{1y} реакции стержня $\vec{R}_1$, приложенные к подвеске, уравновешиваются реакцией опоры, и поэтому вертикальная равнодействующая сила равна нулю. Сила $\mathcal{F}_{1x}$ представляет собой, очевидно, горизонтальную проекцию реакции стержня (связи) R_{1x}. При лагранжевом подходе силы реакции стержня на движение как подвески, так и маятника, т. е. силы $\vec{R}_1$ и $\vec{R}_2$, учитывать нет необходимости (конкретный пример содержится в п. 2). Поэтому из всех действующих сил достаточно принять во внимание только силу тяжести $\vec{\mathcal{F}}_2$, действующую на маятник. Для ее проекций имеем равенства

$$\mathcal{F}_{2x} = 0, \quad \mathcal{F}_{2y} = -m_2 g = -m_2 \frac{\partial \Pi}{\partial y_2},$$

где $\Pi(y_2) = m_2 g y_2$ — потенциальная энергия маятника. В координатах q_1, q_2 $\Pi(y_2)$ выражается формулой

$$V(q_2) = -m_2 l g \cos\alpha.$$

Так как изучаемое движение потенциально, то следует воспользоваться уравнениями Лагранжа (11) § 1, где $j = 1, 2$, $i = 1, 2$, и

$$L = T - V = T_1 + T_2 - V,$$

или, в развернутом виде,

$$L = \frac{m_1 + m_2}{2}\dot{x}_1^2 + \frac{m_2 l}{2}(l\dot{\alpha}^2 + 2\dot{x}_1\dot{\alpha}\cos\alpha) + m_2 l g\cos\alpha. \qquad (1)$$

Из (1) дифференцированием по q_1, $\dot{q}_1$, q_2, $\dot{q}_2$ (напомним, что $q_1 = x_1$, $q_2 = \alpha$) получаем

$$\frac{\partial L}{\partial q_1} = 0, \quad \frac{\partial L}{\partial \dot{q}_1} = (m_1 + m_2)\dot{x}_1 + m_2 l\dot{\alpha}\cos\alpha,$$

$$\frac{\partial L}{\partial q_2} = -m_2 l\sin\alpha(\dot{x}_1\dot{\alpha} + g), \quad \frac{\partial L}{\partial \dot{q}_2} = m_2 l(l\dot{\alpha} + \dot{x}_1\cos\alpha).$$

Подставляя полученные выражения в уравнения Лагранжа и производя дифференцирование по t, приходим к двум уравнениям относительно x_1, α:

$$\begin{aligned} (m_1 + m_2)\ddot{x}_1 + m_2 l\cos\alpha\cdot\ddot{\alpha} &= m_2 l\sin\alpha\cdot\dot{\alpha}^2, \\ \cos\alpha\,\ddot{x}_1 + l\ddot{\alpha} &= -g\sin\alpha, \end{aligned} \qquad (2)$$

представляющим собой модель изучаемой системы. В соответствии с общими свойствами лагранжева формализма уравнения (2) разрешимы относительно $\ddot{x}_1$, $\ddot{\alpha}$ и при известных значениях начальных обобщенных координат и обобщенных скоростей позволяют найти координаты и скорости точек в любой момент времени.

Нелинейная система четвертого порядка (2) легко сводится к уравнению второго порядка, например исключением из нее величины $\ddot{x}_1$:

$$l(m_1 + m_2\sin^2\alpha)\ddot{\alpha} = -\sin\alpha\,[m_2 l\cos\alpha\cdot\dot{\alpha}^2 + (m_1 + m_2)g]. \qquad (3)$$

Этот результат — следствие инвариантности лагранжиана (1) относительно двух однопараметрических семейств преобразований. Первое из них дается формулой $x_1^* = x_1 + \beta$ (L не меняется при сдвиге координаты x_1), а второе — формулой $\alpha^* = \alpha + m\,\mathrm{sign}\,\beta\,2\pi$, где $m = 1, 2, \ldots$; β — параметр преобразования (L не меняется при повороте системы координат на угол, кратный 2π). Согласно теореме Нётер (п. 4 § 1) у системы имеются два первых интеграла, определяемые по формуле (23) § 1, и поэтому ее порядок может быть понижен на две единицы. Еще один интеграл системы очевиден: лагранжиан (1) не зависит явно от времени (консервативность), и сохраняется ее полная энергия $H = T + V$. Это свойство обеспечивает возможность понижения порядка системы (2) еще на одну единицу и сведения (3) к уравнению первого порядка (упр. 1).

Данный пример хорошо иллюстрирует различие между лагранжевым и ньютоновым подходами к описанию движения механических систем. Уравнения Ньютона для подвески и маятника в координатной форме выглядят так:

$$\begin{aligned} m_1\ddot{x}_1 &= R_1(x_2 - x_1)/l, \\ m_2\ddot{x}_2 &= -R_1(x_2 - x_1)/l, \\ m_2\ddot{y}_2 &= -R_2 y_2/l - m_2 g, \end{aligned} \qquad (4)$$

где $R_1 = R_2 = R$ — модуль вектора реакций стержня, приложенных к массам m_1 и m_2 (см. рис. 39), причем очевидно, что $\vec{R}_1 = -\vec{R}_2$. Реакция создается натяжением стержня, который в идеализированной постановке считается абсолютно жестким, и его деформацией пренебрегается.

Три уравнения (4) содержат четыре неизвестных величины: x_1, x_2, y_2, R. Систему (4) можно замкнуть, используя уравнение связи $(x_2 - x_1)^2 + y_2^2 = l^2$, и прийти к некоторому нелинейному уравнению второго порядка (см. также упр. 2). Однако в случае более сложных систем эта громоздкая процедура становится фактически неосуществимой. При составлении уравнений Лагранжа она не требуется (что и послужило первоначальной причиной разработки лагранжева формализма). Кроме того, инвариантные свойства лагранжиана ясно указывают на существование первых интегралов движения, что существенно упрощает исследование.

Получающееся из (3) уравнение первого порядка относительно нетрудно изучить в плоскости функций α, $d\alpha/dt$ (фазовой плоскости) и определить все характеристики движения в зависимости от начальных данных. Ограничимся рассмотрением малых колебаний системы, когда $\alpha \ll 1$. Отбрасывая в (3) члены более высокого порядка малости, приходим к уравнению

$$\ddot{\alpha} = -\frac{g}{l}\frac{m_1 + m_2}{m_1}\alpha,$$

которое, очевидно, имеет общее решение

$$\alpha(t) = A\sin\omega t + B\cos\omega t,$$

где константы A и B определяются из начальных данных, а частота колебаний дается формулой

$$\omega = \sqrt{\frac{g}{l}\left(1 + \frac{m_2}{m_1}\right)}.$$

В сравнении с жестко закрепленным маятником, для которого $\omega_0 = \sqrt{g/l}$, частота увеличивается, зависит от значений m_1, m_2 и растет тем больше, чем больше становится отношение m_2/m_1, что связано со свободным движением точки крепления. Этим же объясняется еще одно отличие, заключающееся в следующем. Пусть в начальный момент $t = 0$ отклонение маятника равно $\alpha(0) > 0$, а его скорость, как и скорость подвески, равна нулю, т. е. энергия системы сосредоточена в потенциальной энергии маятника. Она полностью преобразуется в его кинетическую энергию при прохождении им низшей точки. В этот момент скорость маятника равна $v_{2x} = \dot{x}_1 + l\alpha$. При ее вычислении учтем, что в данном случае $\alpha(t) = \alpha(0)\cos\omega t$ и что $\ddot{x}_1 = -l\ddot{\alpha} - g\alpha$ (последнее равенство вытекает из линеаризованной системы (2)). Таким

образом,

$$\dot{x}_1 = -l\dot{\alpha} - \int_0^t g\alpha(t)\,dt,$$

$$v_{2x} = -g\int_0^t \alpha(0)\cos\omega t\,dt.$$

интересующий нас момент $t = \pi/(2\omega)$

$$v_{2x}(\alpha = 0) = -g\alpha(0)\int_0^{\pi/(2\omega)} \cos\omega t\,dt = -\frac{g\alpha(0)}{\omega}.$$

Это значение в ω/ω_0 раз меньше максимальной скорости маятника на жесткой подвеске — запасенная вначале энергия частично переходит в кинетическую энергию подвески.

Если же $m_1 \to \infty$ (очень массивная подвеска), то, естественно, как малые, так и конечные колебания системы совпадают с движением жестко закрепленного маятника.

2. Непотенциальные колебания. Учтем теперь действие сил трения на маятник и подвеску, считая их пропорциональными скоростям:

$$\vec{F}_1 = -\mu_1\vec{v}_1, \qquad \vec{F}_2 = -\mu_2\vec{v}_2, \qquad \mu_1 > 0, \qquad \mu_2 > 0.$$

Так как силы трения зависят от скоростей, то движение непотенциально, и следует использовать уравнения Лагранжа (9) § 1 с обобщенными силами в правых частях. Выбирая, как и прежде, $q_1 = x_1$, $q_2 = \alpha$, получаем из общей формулы (10) § 1

$$\Phi_1 = \mathcal{F}_{1x}\frac{\partial x_1}{\partial q_1} + \mathcal{F}_{1y}\frac{\partial y_1}{\partial q_1} + \mathcal{F}_{2x}\frac{\partial x_2}{\partial q_1} + \mathcal{F}_{2y}\frac{\partial y_2}{\partial q_1},$$

$$\Phi_2 = \mathcal{F}_{1x}\frac{\partial x_1}{\partial q_2} + \mathcal{F}_{1y}\frac{\partial y_1}{\partial q_2} + \mathcal{F}_{2x}\frac{\partial x_2}{\partial q_2} + \mathcal{F}_{2y}\frac{\partial y_2}{\partial q_2},$$

где $\mathcal{F}_{1x}$, $\mathcal{F}_{1y}$, $\mathcal{F}_{2x}$, $\mathcal{F}_{2y}$ — компоненты равнодействующих сил, приложенных к массам m_1, m_2 (см. рис. 39). Приняв во внимание, что $\mathcal{F}_{1y} = 0$, и равенства $\partial x_1/\partial q_1 = \partial x_2/\partial q_1 = 1$, $\partial y_1/\partial q_1 = \partial x_1/\partial q_2 = 0$, $\partial x_2/\partial q_2 = l\cos q_2$, $\partial y_2/\partial q_2 = l\sin q_2$, упростим выражения для Φ_1, Φ_2:

$$\begin{aligned} \Phi_1 &= \mathcal{F}_{1x} + \mathcal{F}_{2x}, \\ \Phi_2 &= \mathcal{F}_{2x}l\cos q_2 + \mathcal{F}_{2y}l\sin q_2. \end{aligned} \tag{5}$$

Выразим компоненты действующих сил в координатах $q_1 = x_1$, $q_2 = \alpha$:

$$\mathcal{F}_{1x} = F_{1x} + R_{1x} = -\mu_1 v_{1x} + R_{1x} = -\mu_1 \dot{x}_1 + R\sin\alpha,$$

$$\mathcal{F}_{2x} = F_{2x} + R_{2x} = -\mu_2 v_{2x} + R_{2x} = -\mu_2(\dot{x}_1 + l\dot{\alpha}\cos\alpha) - R\sin\alpha,$$

$$\mathcal{F}_{2y} = F_{2y} + R_{2y} - mg = -\mu_2 v_{2y} + R_{2y} - m_2 g =$$

$$= -\mu_2 l\dot{\alpha}\sin\alpha + R\cos\alpha - m_2 g,$$

где R — величина силы реакции стержня. Подставляя их в (5), получим

$$\begin{aligned} \Phi_1 &= -(\mu_1 + \mu_2)\dot{x}_1 - \mu_2 l\dot{\alpha}\cos\alpha, \\ \Phi_2 &= \mu_2 l\dot{x}_1\cos\alpha - \mu_2 l^2\dot{\alpha} - m_2 gl\sin\alpha. \end{aligned} \tag{6}$$

В соответствии с общим свойством лагранжева формализма реакции связей, как видно из (6), не вошли в окончательное выражение для Φ_1, Φ_2. Кинетическая энергия системы T найдена в п. 1:

$$T = \frac{m_1 + m_2}{2}\dot{x}_1^2 + \frac{m_2 l}{2}(l\dot{\alpha}^2 + 2\dot{x}_1\dot{\alpha}\cos\alpha).$$

Вычисляя аналогично п. 1 производные $dT/dq_{1,2}$ и $dT/d\dot{q}_{1,2}$ и дифференцируя по t, придем к уравнениям Лагранжа применительно к рассматриваемой системе

$$\begin{aligned} (m_1 + m_2)\ddot{x}_1 + m_2 l\cos\alpha\cdot\ddot{\alpha} &= m_2 l\sin\alpha\cdot\dot{\alpha}^2 + \Phi_1, \\ m_2 l\cos\alpha\cdot\ddot{x}_1 + m_2 l^2\ddot{\alpha} &= \Phi_2. \end{aligned} \tag{7}$$

Как и для потенциального движения, нелинейная система четвертого порядка (7) разрешается относительно старших производных, и при заданных начальных значениях $x_1(0)$, $\dot{x}_1(0)$, $\alpha(0)$, $\dot{\alpha}(0)$ из нее определяются положения и скорости масс m_1, m_2 в любой момент времени.

Однако, в отличие от системы (2), у изучаемого движения нет трех первых интегралов (теорема Нётер справедлива для потенциальных движений), и ее порядок может быть понижен лишь на единицу (упр. 3). Еще одно отличие заключается в виде балансного энергетического соотношения

$$E(0) = E(t) + A(t), \tag{8}$$

где $E(0)$ — полная начальная энергия системы, $E(t) = T(t) + V(t)$ — текущая полная энергия, $A(t)$ — работа сил трения, произведенная к моменту t. Механический смысл (8) состоит в том, что величина утраченной (диссипированной) энергии системы равна работе, совершенной над нею непотенциальными силами трения.

Получим равенство (8) для простоты в случае $\mu_2 = 0$ (трение испытывает лишь подвеска). Из (6) имеем

$$\Phi_1 = -\mu_1\dot{x}_1, \quad \Phi_2 = -m_2 lg\sin\alpha.$$

Подставим эти выражения в (7), умножим первое уравнение на $\dot{x}_1$, второе — на $\dot{\alpha}$, сложим оба уравнения и придем к равенству

$$\begin{aligned} \ddot{x}_1((m_1 + m_2)\dot{x}_1 + m_2 l\cos\alpha\cdot\dot{\alpha}) + \ddot{\alpha}(m_2 l\dot{x}_1\cos\alpha + m_2 l^2\dot{\alpha}) &= \\ = m_2 l\sin\alpha\cdot\dot{\alpha}\cdot\dot{x}_1 - m_2 lg\sin\alpha\cdot\dot{\alpha} - \mu_1\dot{x}_1^2. \end{aligned} \tag{9}$$

Соотношение (9) совпадает с продифференцированным по времени равенством (8), в котором полная энергия в момент t есть

$$E(t) = \frac{m_1 + m_2}{2}\dot{x}_1^2 + \frac{m_2 l}{2}(l\dot{\alpha}^2 + 2\dot{x}_1\dot{\alpha}\cos\alpha) - m_2 lg\cos\alpha,$$

а работа силы трения дается формулой

$$A(t) = -\int_0^t F_1\,dx_1 = \int_0^t \mu_1 v_1\,dx_1 = \int_0^t \mu_1\frac{dx_1}{dt}\,dx_1 = \int_0^t \mu_1\left(\frac{dx_1}{dt}\right)^2 dt.$$

Таким образом, (9) эквивалентно равенству

$$\frac{d}{dt}(E(t) + A(t)) = \frac{d}{dt}\left(E(t) + \int_0^t \mu_1\left(\frac{dx_1}{dt}\right)^2 dt\right) = 0,$$

или, что то же самое, равенству (8). Так как

$$\frac{dE(t)}{dt} = -\mu_1\left(\frac{dx_1}{dt}\right)^2 \leqslant 0,$$

то в сравнении с консервативной системой п. 1 полная энергия в данном случае не постоянна, а убывает со временем.

При малых колебаниях жестко закрепленного маятника из (7) получаем уравнение

$$\ddot{\alpha} = -\frac{\mu_2}{m_2}\dot{\alpha} - \frac{g}{l}\alpha,$$

эквивалентное уравнению движения в вязкой среде шарика на пружинке (п. 3 § 4 гл. I) и имеющее простое общее решение.

3. Малые колебания струны. Применимость принципа Гамильтона не ограничивается системами материальных точек. Он распространяется на объекты, не являющиеся, строго говоря, совокупностями точечных масс. Примером служит упругая нить или струна — сплошная среда, которую, однако, можно рассматривать как множество примыкающих друг к другу материальных точек. Будем считать, что толщина струны много меньше ее длины l и что она имеет постоянную линейную плотность ρ_0. Натянутая с силой F_0 струна в состоянии равновесия неподвижна и представляет собой прямую линию. При отклонении от равновесия, например, в результате удара, стру-

на изгибается, ее участки начинают двигаться (рис. 40). Колебания считаются плоскими и малыми — их амплитуда значительно меньше длины струны. Это предположение позволяет пренебречь продольными смещениями и скоростями участков струны, рассматривая только поперечное их движение.

Представим струну как совокупность N материальных точек с равными массами $m_i = \rho_0 l/N = \rho_0 \Delta x$, $i = 1, \ldots, N$. Длина каждого участка струны, содержащего массу m_i, равная Δx в положении равновесия и несильно, ввиду малости колебаний, изменяющаяся при отклонении от него, считается малой. Поэтому положение i-й «материальной точки» в любой момент времени можно охарактеризовать величинами $x_i(t)$ (продольной координатой центра i-го отрезка) и $y_i(t)$ (поперечным отклонением центра отрезка от положения равновесия).

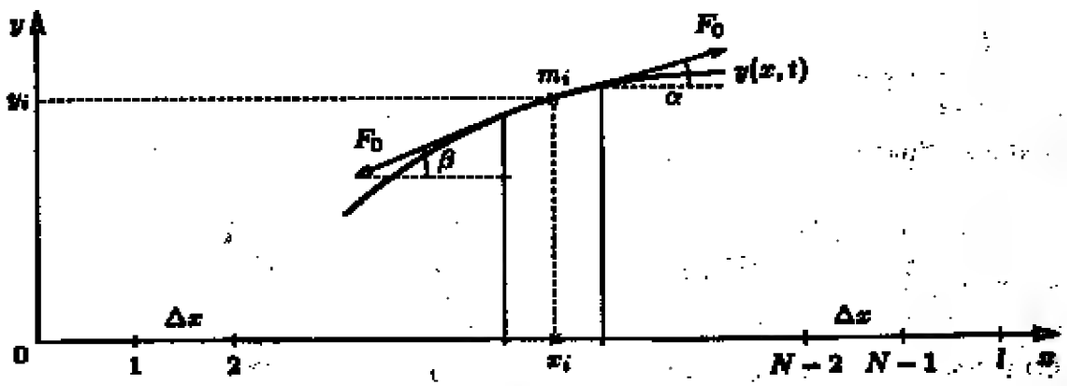


Рис. 40

Введенные таким образом обобщенные (в данном случае — декартовы) координаты полностью описывают плоское движение рассматриваемой системы. В силу малости отклонений, как уже отмечалось, $dx_i(t)/dt = v_{ix} = 0$, т. е. координаты x_i не зависят от времени.

Составляющие струну «материальные точки» связаны между собой и при $N \to \infty$, $\Delta x \to 0$ образуют для любого момента времени в плоскости x, y некоторую кривую

$$y = y(x, t). \tag{10}$$

В отличие от рассматривавшихся выше механических связей, которые всегда считаются заданными, «связь» (10) неизвестна, и функция $y(x,t)$ подлежит определению. Если она найдена, то тем самым известны координаты $y_i = y(x_i, t)$, и поскольку координаты x_i не меняются с течением времени, полностью известно движение системы.

Кинетическая энергия i-й массы определяется по формуле

$$T_i = \frac{1}{2} m_i v_{iy}^2 = \frac{1}{2} m_i \left(\frac{dy_i}{dt}\right)^2 = \frac{1}{2} m_i \left(\frac{\partial y_i}{\partial t}\right)^2.$$

При получении этой формулы использовано равенство $dy_i/dt = \partial y_i/\partial t$, справедливое потому, что $dx_i/dt = 0$. Полная кинетическая энергия

системы равна

$$T = \frac{1}{2} \sum_{i=1}^{N} m_i \left(\frac{\partial y_i}{\partial t}\right)^2 = \frac{1}{2} \sum_{i=1}^{N} \rho_0 \left(\frac{\partial y}{\partial t}\right)_i^2 \Delta x. \tag{11}$$

Вычислим теперь силы, действующие на i-ю массу. По предположению о малых колебаниях $\mathcal{F}_{ix} = 0$. Вертикальная компонента $\mathcal{F}_{iy}$ определяется как сумма вертикальных компонент силы натяжения струны на правом и левом концах i-го отрезка. Его удлинение при отклонении от положения равновесия мало, поэтому силу натяжения струны можно считать неизменной и равной F_0. Тогда вертикальные составляющие сил, приложенных к правому и левому концам, соответственно равны

$$\mathcal{F}_- = -F_0 \sin\alpha = -F_0 \frac{\partial y}{\partial x}\left(x_i + \frac{\Delta x}{2}\right),$$

$$\mathcal{F}_+ = F_0 \sin\beta = F_0 \frac{\partial y}{\partial x}\left(x_i - \frac{\Delta x}{2}\right)$$

(см. рис. 40; напомним, что сила натяжения направлена по касательной к струне). В итоге получаем

$$\mathcal{F}_{iy} = \mathcal{F}_+ + \mathcal{F}_- = F_0 \left[y_x(x - \Delta x/2) - y_x(x + \Delta x/2)\right],$$

или, учитывая малость отрезка Δx,

$$\mathcal{F}_{iy} = -F_0 (y_{xx})_i \Delta x.$$

Примем во внимание, что из (10) следует $\partial y = y_x \cdot \partial x$, и перепишем последнее выражение в виде

$$\mathcal{F}_{iy} = -F_0 \Delta x \left(\frac{\partial}{\partial x} y_x\right)_i = -F_0 \Delta x \left(y_x \frac{\partial}{\partial y} y_x\right)_i = -\frac{1}{2} F_0 \Delta x \left(\frac{\partial}{\partial y} y_x^2\right)_i.$$

Отсюда сразу видно, что все силы $\mathcal{F}_{iy}$, $i = 1, \ldots, N$, потенциальны, причем потенциальная энергия i-й массы дается формулой

$$V_i = \frac{1}{2} F_0 (y_x^2)_i \Delta x,$$

а полная потенциальная энергия струны есть

$$V = \frac{1}{2} F_0 \sum_{i=1}^{N} (y_x^2)_i \Delta x. \tag{12}$$

Поскольку движение системы потенциально, то к нему применим принцип Гамильтона, сформулированный в § 1. Из (11), (12) получаем лагранжиан

$$L(y, \dot{y}, t) = T - V = \frac{1}{2} \sum_{i=1}^{N} \left[\rho_0 \left(\frac{\partial y}{\partial t}\right)^2 - F_0 \left(\frac{\partial y}{\partial x}\right)^2\right]_i \cdot \Delta x, \qquad \dot{y} = \frac{\partial y}{\partial t}.$$

Действие по Гамильтону вычисляется по формуле

$$Q(y,\dot y,t)=\int\limits_{t_0}^{t_1} L\,dt=\int\limits_{t_0}^{t_1}\frac{1}{2}\sum_{i=1}^{N}\left[\rho_0\left(\frac{\partial y}{\partial t}\right)^2-F_0\left(\frac{\partial y}{\partial x}\right)^2\right]_i\Delta x\,dt,\qquad(13)$$

где t_0, t_1 — два произвольных момента времени, в которые система имеет координаты $y_i(t_0)$, $y_i(t_1)$. Перейти из состояния с $y_i(t_0)$ в состояние с $y_i(t_1)$ можно, вообще говоря, многими путями. Принцип Гамильтона выделяет из этого множества «истинный» прямой путь, для которого вариация действия δQ равна нулю. Вычислим с помощью (13) вариацию δQ, придавая вариацию δy_i координатам и $\delta\dot y_i$ — скоростям i-х точек:

$$\delta Q=\delta\int\limits_{t_0}^{t_1} L\,dt=$$

$$=\frac{1}{2}\int\limits_{t_0}^{t_1}\sum_{i=1}^{N}\left[2\rho_0\frac{\partial y}{\partial t}\delta\left(\frac{\partial y}{\partial t}\right)-F_0\frac{\partial}{\partial y}\left[\left(\frac{\partial y}{\partial x}\right)^2\right]\cdot\delta y\right]_i\Delta x\,dt.\qquad(14)$$

Первый член в правой части (14) возникает в результате варьирования скорости $\dot y_i$, а второй — при варьировании координаты y_i. Для преобразования первого члена проинтегрируем его по частям, используя перестановочность операций варьирования и дифференцирования по t, и, принимая во внимание, что $\delta y(t_0)=\delta y(t_1)=0$, получим

$$\frac{1}{2}\int\limits_{t_0}^{t_1}\sum_{i=1}^{N}\left[2\rho_0\frac{\partial y}{\partial t}\delta\left(\frac{\partial y}{\partial t}\right)\right]_i\Delta x\,dt=\int\limits_{t_0}^{t_1}\sum_{i=1}^{N}\rho_0\left(\frac{\partial^2 y}{\partial t^2}\right)_i\delta y_i\,\Delta x\,dt.$$

Второй член с учетом равенства $\partial y=y_x\partial x$ переписывается в виде

$$-\frac{1}{2}\int\limits_{t_0}^{t_1}\sum_{i=1}^{N}F_0\frac{\partial}{\partial y}\left[\left(\frac{\partial y}{\partial x}\right)^2_i\right]\delta y_i=-\int\limits_{t_0}^{t_1}F_0\left(\frac{\partial^2 y}{\partial x^2}\right)_i\delta y_i\,\Delta x\,dt.$$

После подстановки этих выражений в (14) получим

$$\delta Q=\int\limits_{t_0}^{t_1}\sum_{i=1}^{N}\left(\rho_0\frac{\partial^2 y}{\partial t^2}-F_0\frac{\partial^2 y}{\partial x^2}\right)_i\delta y_i\,\Delta x\,dt.$$

Перейдем от дискретной системы «материальных точек», которой первоначально была заменена струна, к сплошной среде. Для этого в последнем равенстве устремим $N\to\infty$, заменим Δx на dx и опустим индекс i:

$$\delta Q=\int\limits_{t_0}^{t_1}\int\limits_{0}^{l}\left(\rho_0\frac{\partial^2 y}{\partial t^2}-F_0\frac{\partial^2 y}{\partial x^2}\right)\delta y\,dx\,dt.$$

На прямом пути $\delta Q=0$, что возможно, лишь если в последнем равенстве подынтегральное выражение равно нулю при всех x и t, т. е.

$$\frac{\partial^2 y}{\partial t^2}=a_0^2\frac{\partial^2 y}{\partial x^2},\qquad a_0^2=\frac{F_0}{\rho_0},\quad 0<x<l,\quad t>0.\qquad(15)$$

Следовательно, при малых колебаниях струны ее отклонение подчиняется уравнению (15), из которого при соответствующих краевых условиях определяется функция $y=y(x,t)$.

Принцип Гамильтона применительно к рассматриваемой ситуации можно трактовать как способ получения уравнения «связи» (10). В силу свойств лагранжиана полная энергия $H=T+V$ струны сохраняется со временем (консервативность движения нетрудно установить также непосредственно из уравнения (15); см. упр. 4).

Уравнение малых колебаний струны (15) ($\omega=2\pi a_0/\lambda$ — частота колебаний с длиной волны λ) — линейное уравнение второго порядка в частных производных гиперболического типа. Принцип суперпозиции позволяет получить его общее решение как сумму частных решений, используя соответствующие методы теории уравнений математической физики.

Основной краевой задачей для (15) является первая краевая задача на отрезке $[0,l]$, когда для однозначного определения решения задаются начальные отклонения $y(x,0)=y_0(x)$, $0<x<l$, скорости $\dot y(x,0)=\dot y_0(x)$, $0<x<l$, и граничные значения функции $y(0,t)=y_1(t)$, $t>0$ и $y(0,l)=y_2(t)$, $t>0$. Основная задача допускает различные модификации, самая простая из которых — задача Коши, решаемая на всей прямой $-\infty<x<\infty$. Такая идеализация оправдана в случае, если рассматривается движение центральной части струны в течение непродолжительного времени, и влиянием границ можно пренебречь. Для решения задачи Коши достаточно знать начальные скорости и координаты струны, т. е. функции $y_0(x)$, $\dot y_0(x)$ при $-\infty<x<\infty$.

Для частного вида движений, обладающих свойством $\partial y/\partial t=c\,\partial y/\partial x$, $c=\mathrm{const}$ (простые волны), уравнение (15) переходит в изученное в § 1 гл. II гиперболическое уравнение первого порядка, или уравнение переноса (см. также упр. 5).

Отметим, что обычно уравнение (15) получают с помощью непосредственного применения второго закона Ньютона и закона Гука к элементарному участку струны. При этом предположения о малости колебаний, однородности струны и т. д. в обоих подходах одни и те же. Поэтому и математические модели движения струны в обоих случаях одинаковы.

4. Электромеханическая аналогия. Распространение принципа Гамильтона возможно не только на процессы движения сплошных сред, но и на некоторые немеханические объекты. Рассмотрим уже изучавшийся в § 5 гл. I колебательный контур, состоящий из конденсатора с емкостью C_0 и катушки с индуктивностью L_0. В начальный момент времени цепь разомкнута, заряд сосредоточен на обкладках конденсатора. При замыкании цепи конденсатор начинает разряжаться, и по ней идет ток.

Электромеханическая аналогия заключается в следующем. Обобщенной координате отвечает заряд на обкладках конденсатора — неизвестная функция времени $q = q(t)$. Величина электрического тока $\dot{q}(t) = dq(t)/dt = i(t)$ играет роль обобщенной скорости. Для правильного определения аналогов кинетической энергии (энергии движения) и потенциальной энергии (энергии конденсатора) будем руководствоваться следующими наводящими соображениями. Энергия перемещающихся по проводнику зарядов (энергия тока) пропорциональна квадрату скорости v их направленного движения. С другой стороны, заряд, проходящий через поперечное сечение S проводника в единицу времени (ток), равен $i = q_0 nSv$, где q_0, n — величины элементарного заряда и объемной концентрации переносчиков тока соответственно. Следовательно, энергия движения частиц $T \sim v^2 \sim i^2$, т. е. пропорциональна квадрату тока $\dot{q}(t) = i(t)$. Коэффициент пропорциональности (аналог массы) берется равным L_0, т. е.

$$T = T(\dot{q}) = \frac{1}{2} L_0 \dot{q}^2.$$

Потенциальная энергия контура содержится в конденсаторе. Для его зарядки требуется затратить определенную работу по разделению разноименных зарядов. Согласно закону Кулона сила, препятствующая этому, как функция зарядов q_1, q_2 пропорциональна их произведению $q_1 q_2$ (если $q_1 = q_2 = q$, то сила пропорциональна q^2). Таким образом, работа по разделению зарядов, т. е. потенциальная энергия системы V, пропорциональна квадрату «обобщенной координаты»:

$$V = V(q) = \frac{1}{2C_0} q^2,$$

где $1/C_0$ — аналог коэффициента упругой силы в законе Гука (система шарик—пружина), или величине $\sqrt{g/l}$ в случае колебаний маятника.

Примем теперь во внимание, что действующие в контуре силы имеют чисто электростатическое происхождение (сопротивлением проводников пренебрегается, т. е. «трение» отсутствует, как отсутствуют и потери энергии на излучение электромагнитных волн). По закону Кулона эти силы определяются «обобщенной координатой» q и не зависят от $\dot{q}$. В этом смысле силы «потенциальны», и вместе с ними «потенциальна» рассматриваемая система. Поэтому у нее существует «лагранжиан» $L = T - V$ и к ней применим аналог принципа Гамильтона: для «истинного» пути системы вариация «действия»

$$Q = \int_{t_0}^{t_1} L\,dt$$

равна нулю (здесь, как обычно, t_0, t_1 — два произвольно взятых момента времени). Пусть функция $q(t, 0) = q^0(t)$ отвечает прямому пути системы в промежутке $t_0 < t < t_1$. Вариация координаты $q(t, \alpha)$, $\alpha \neq 0$, равна $\delta q = q(t, \alpha) - q^0(t)$, где $q(t, \alpha)$ — все возможные «траектории»,

имеющие одинаковые координаты $q(t_0, \alpha)$, $q(t_1, \alpha)$. Для вариации «действия» имеем

$$\delta Q = \delta \int_{t_0}^{t_1} L\,dt = \delta \int_{t_0}^{t_1} \frac{1}{2}\left(L_0 \dot{q}^2 - \frac{1}{C_0} q\right) dt = \frac{1}{2} \int_{t_0}^{t_1} [L(q) - L(q^0)]\,dt.$$

Так как $q = q^0 + \delta q$, то подынтегральное выражение можно представить в виде

$$L(q) - L(q^0) = L_0\left[(\dot{q}^0)^2 + 2\dot{q}^0\,\delta\dot{q} + \delta\dot{q}^2\right] - \\ - \frac{1}{C_0}\left[(q^0)^2 + 2q^0\,\delta q + \delta q^2\right] - L_0(\dot{q}^0)^2 + \frac{1}{C_0}(q^0)^2.$$

Отбросив в нем члены второго порядка малости, получим

$$\delta Q = \int_{t_0}^{t_1} \left(L_0 \dot{q}^0\,\delta\dot{q} - \frac{1}{C_0} q^0\,\delta q\right) dt.$$

Интегрируя член $L_0 q^0\,\delta\dot{q}$, где $\delta\dot{q} = d(\delta q)/dt$, по частям и учитывая, что $\delta q(t_0) = \delta q(t_1) = 0$, придем к окончательному выражению для δQ

$$\delta Q = \int_{t_0}^{t_1} \left(L_0 \ddot{q}^0 + \frac{1}{C_0} q^0\right) \delta q\,dt = 0.$$

Из него для заряда $q(t)$ (верхний индекс у q^0 опущен) следует уравнение

$$L_0 \ddot{q} = -\frac{1}{C_0} q,$$

описывающее колебания в электрическом контуре и полученное в § 5 гл. I другим способом. Заметим, что полная энергия колебаний $H = T + V$ сохраняется со временем, что согласуется с инвариантностью «лагранжиана» относительно сдвига по времени.

Рассмотренная аналогия применима также к электрическим цепям гораздо более сложной конфигурации, на ее основе строятся математические модели протекающих в них процессов. Приведенный пример далеко не единственная иллюстрация широкой применимости принципа Гамильтона и других вариационных принципов. Они часто используются при построении математических моделей не только механических или физических, но и многих химических, биологических и иных явлений.

УПРАЖНЕНИЯ

1. Используя замену $d\alpha/dt = v$, сведите уравнение (3) к виду $dv/d\alpha = f(v, \alpha)$.
2. Переходом в уравнениях (4) к координатам $x_2 = x_1 + l\sin\alpha$, $y_2 = -l\cos\alpha$ получите уравнение (3).

3. В систему (7) величина x_1 явным образом не входит. Покажите, используя это свойство, что (7) сводится к системе уравнений третьего порядка относительно функций $X(t) = dx_1/dt$, $Y(t) = d\alpha/dt$, $Z(t) = \alpha(t)$.

4. Умножьте обе части уравнения (15) на $\partial y/\partial t$ и, интегрируя полученное равенство от $t = 0$ до $t > 0$ и от $x = 0$ до $x = l$, убедитесь в том, что $H(t) = H(0)$ для любых $t > 0$.

5. Получите решение уравнения (16), описывающее движения струны, для которых $\partial y/\partial t = c\partial y/\partial x$.

§ 3. Уравнение Больцмана и производные от него

Построим иерархическую цепочку моделей, описывающих динамику большого числа материальных частиц. Основой первоначальной модели будет служить кинетическое уравнение Больцмана для функции распределения. Из него по принципу «сверху вниз» путем последовательного введения упрощающих предположений получим модели для вязкого теплопроводного газа, уравнения Эйлера и уравнения акустики.

1. Описание совокупности частиц с помощью функции распределения. В пространстве x, y, z находится «жидкость» или «газ» — большое число материальных частиц (молекул, атомов, электронов, ионов), свободно движущихся в промежутках между взаимными «столкновениями». Если известны координаты и скорости любой частицы как функции времени, то тем самым полностью известны основные характеристики рассматриваемого ансамбля. Они определяются начальным состоянием каждой частицы, характером взаимодействия между ними (например, электроны и ионы притягиваются или отталкиваются по закону Кулона), а также действующими на них внешними силами.

Задавая в момент $t = 0$ положение и скорости частиц и зная все действующие на них силы, можно в принципе решить основную задачу механики (§ 1) для рассматриваемой системы (имеется в виду, что они подчиняются законам классической механики). Такой подход (применительно к обычному газу он называется *молекулярной динамикой*) дает исчерпывающую информацию и занимает высшее место в иерархии математических описаний газа. Однако при достаточно большом количестве частиц реализовать его практически невозможно в том числе и потому, что их начальные положения и скорости никогда в точности не известны. Кроме того, как правило, нет необходимости прослеживать поведение каждой частицы, поскольку интерес представляют лишь средние характеристики системы такие, как плотность, скорость, температура и т. п.

Поэтому описание с помощью *первопринципов*, т. е. с применением второго закона Ньютона к каждой из частиц, используется только в специальных случаях. Переход к следующему звену в иерархии моделей основан на отказе от изучения индивидуальной судьбы частиц. Вводится статистическое вероятностное описание их ансамбля с помощью *функции распределения* $f(\vec{r}, \vec{v}, t)$, где $\vec{r}$, $\vec{v}$ — радиус-вектор и скорость соответственно. Функция f зависит от «координат» шестимерного *фазового пространства* x, y, z, v_x, v_y, v_z (пространства

состояний) и времени (ср. с функцией распределения фотонов в § 3 гл. II). Величина

$$f(\vec{r}, \vec{v}, t)\, d\vec{r}\, d\vec{v}$$

по определению равна числу частиц, находящихся в момент t (точнее, их среднему значению за короткий промежуток времени dt) в элементе $d\vec{r}\,d\vec{v}$ фазового пространства, т. е. имеющих координаты в интервале от $\vec{r}$ до $\vec{r} + d\vec{r}$ и скорости в диапазоне от $\vec{v}$ до $\vec{v} + d\vec{v}$. Элемент $d\vec{r}\,d\vec{v}$ считается малым в сравнении с характерными размерами системы, но содержащим достаточно большое число частиц. Для простоты рассматривается газ, состоящий из частиц одного сорта.

Через функцию распределения вычисляются средние величины, характеризующие состояние газа в пространстве и во времени. Например, очевидно, что выражение

$$n(\vec{r}, t) = \int f(\vec{r}, \vec{v}, t)\, d\vec{v},$$

где интегрирование ведется по всем скоростям, есть не что иное, как число частиц, находящихся в единице физического объема с координатой $\vec{r}$ в момент времени t.

В общем случае средние величины находятся следующим образом. Пусть $\Phi(\vec{v})$ — произвольная функция от скорости частицы (кинетическая энергия, скорость и т. д.). Обозначим через $\sum \Phi$ усредненную за время dt сумму значений функции Φ по всем частицам, находящимся в элементарном физическом объеме $d\vec{r}$. Тогда среднее, т. е. приходящееся на одну частицу, значение Φ получается делением $\sum \Phi$ на число частиц в объеме $d\vec{r}$, которое равно $n\, d\vec{r}$:

$$\langle \Phi \rangle = \frac{\sum \Phi(\vec{v})}{n\, d\vec{r}}.$$

С другой стороны, число частиц в элементарном фазовом объеме $d\vec{r}\,d\vec{v}$ равно $f\, d\vec{r}\,d\vec{v}$, и каждая из них вносит в $\sum \Phi$ вклад, равный $\Phi(\vec{v})$, а их общий элементарный вклад равен $\Phi f\, d\vec{r}\,d\vec{v}$. Учтем теперь, что частицы, находящиеся в физическом объеме $d\vec{r}$, могут иметь любые скорости $\vec{v}$. Поэтому для получения их полного вклада необходимо просуммировать элементарные вклады по всем скоростям:

$$\sum \Phi(\vec{v}) = d\vec{r} \int \Phi(\vec{v})\, f(\vec{r}, \vec{v}, t)\, d\vec{v}.$$

Сравнивая две последние формулы, получим для среднего значения Φ

$$\langle \Phi \rangle = \frac{\int \Phi f\, d\vec{v}}{\int f\, d\vec{v}} = \frac{\int \Phi f\, d\vec{v}}{n}. \qquad (1)$$

Такова связь любой средней функции, характеризующей газ, с функцией распределения. Например, если $\Phi(\vec{v}) = \vec{v}$, то для скорости газа имеем

$$\vec{V}(\vec{r},t) = \int \vec{v} f(\vec{r},\vec{v},t)\, d\vec{v} \cdot n^{-1}.$$

Аналогично, по известной функции распределения можно вычислить как функции $\vec{r}$ и t и другие макроскопические величины, описывающие состояние газа.

2. Уравнение Больцмана для функции распределения. Дадим нестрогий вывод этого уравнения, основываясь на следующих предположениях:

1) время столкновения (непосредственного эффективного силового взаимодействия) частиц много меньше времени, проходящего между их столкновениями;

2) влиянием внешних сил, действующих на частицы (гравитационных, электрических и т. д.), можно пренебречь;

3) частицы не расщепляются и не объединяются.

Рассмотрим частицы, находящиеся в момент t в фазовом объеме $d\vec{r}\,d\vec{v}$ (имеющие координаты и скорости в диапазонах от $\vec{r}$ до $\vec{r}+d\vec{r}$ и от $\vec{v}$ до $\vec{v}+d\vec{v}$ соответственно). Пусть столкновения между ними отсутствуют. Тогда к моменту $t^* = t + dt$ скорости частиц не изменяются, а их координаты получат приращения, соответствующие начальным скоростям:

$$\vec{v}^* = \vec{v}, \quad \vec{r}^* = \vec{r} + \vec{v}\,dt,$$

где, напомним, величины $\vec{v}$ и $\vec{r}$ лежат в указанных выше диапазонах. Вычислим фазовый объем частиц $d\vec{r}^*\,d\vec{v}^*$ в момент t^*. Из последних двух равенств имеем

$$d\vec{v}^* = d\vec{v}, \quad d\vec{r}^* = d\vec{r} + d\vec{v}\,dt = d\vec{r} + \frac{d\vec{v}}{dt}(dt)^2 + \ldots,$$

т. е. $d\vec{r}\,d\vec{v} = d\vec{r}^*\,d\vec{v}^*$, и фазовый объем частиц сохраняется с точностью до членов порядка $(dt)^2$. При этом число частиц $f(\vec{r},\vec{v},t)\,d\vec{r}\,d\vec{v}$, находившихся в объеме $d\vec{r}\,d\vec{v}$, равно их числу $f(\vec{r}^*,\vec{v}^*,t^*)\,d\vec{r}^*\,d\vec{v}^*$ в объеме $d\vec{r}^*\,d\vec{v}^*$. Другими словами, в отсутствие столкновений $f = f^*$, т. е. функция распределения не меняется со временем, частицы лишь сменяют фазовый объем (в данном случае только свои координаты).

Учтем теперь возможность столкновений, введя понятие *интеграла столкновений* $S(f)$. По своему смыслу величина $S(f)\,d\vec{r}\,d\vec{v}\,dt$ — разность между числом частиц, вышедших благодаря столкновениям за время dt из объема $d\vec{r}\,d\vec{v}$ (и не попавших в объем $d\vec{r}^*\,d\vec{v}^*$), и числом частиц, попавших из-за столкновений за время dt в объем $d\vec{r}^*\,d\vec{v}^*$ (и не находившихся в первоначальном объеме $d\vec{r}\,d\vec{v}$). Тогда уравнение баланса (сохранения) числа частиц за время dt при переходе от объема $d\vec{r}\,d\vec{v}$ к объему $d\vec{r}^*\,d\vec{v}^*$ записывается следующим образом:

$$f(\vec{r}^*,\vec{v}^*,t^*)\,d\vec{r}^*\,d\vec{v}^* - f(\vec{r},\vec{v},t)\,d\vec{r}\,d\vec{v} = -S(f)\,d\vec{r}\,d\vec{v}\,dt.$$

Очевидно, что (в отличие от процесса без столкновений) при учете столкновений, вообще говоря, $f \neq f^*$.

Разложим левую часть уравнения баланса в ряд Тейлора, опуская члены порядка $(dt)^2$ и выше:

$$f(\vec{r},\vec{v},t)\,d\vec{r}^*\,d\vec{v}^* + \frac{\partial f}{\partial t}\,d\vec{r}^*\,d\vec{v}^*\,dt + \frac{\partial f}{\partial \vec{r}}\frac{d\vec{r}}{dt}\,d\vec{r}^*\,d\vec{v}^*\,dt +$$
$$+ \frac{\partial f}{\partial \vec{v}}\frac{d\vec{v}}{dt}\,d\vec{r}^*\,d\vec{v}^*\,dt - f(\vec{r},\vec{v},t)\,d\vec{r}\,d\vec{v} = -S(f)\,d\vec{r}\,d\vec{v}\,dt.$$

В полученном равенстве $d\vec{r}/dt = \vec{v}$ и $m\,d\vec{v}/dt = \vec{F}$, где $\vec{F}$ — внешняя сила, которая по предположению равна нулю, m — масса частицы. Выберем промежуток dt настолько малым, что за время от t до $t^* = t + dt$ частицы не сталкиваются. Тогда фазовый объем сохраняется, и $d\vec{r}\,d\vec{v} = d\vec{r}^*\,d\vec{v}^*$. Поделив обе части равенства на $d\vec{r}\,d\vec{v}\,dt$ и устремив dt к нулю, придем к *уравнению Больцмана*

$$\frac{\partial f}{\partial t} + \vec{v}\frac{\partial f}{\partial \vec{r}} + S(f) = 0, \tag{2}$$

которое является следующей по сложности моделью (после моделей, получаемых из первопринципов) в иерархии математических описаний газа. Оно представляет собой нелинейное интегро-дифференциальное уравнение. Его конкретная запись зависит от характера столкновений частиц, т. е. от вида функции $S(f)$. Заметим, что (2) легко обобщается на газ, состоящий из частиц разного сорта, и на случай, когда внешняя сила $\vec{F}$ отлична от нуля (она может быть вызвана, например, наличием электромагнитных полей). При отсутствии столкновений, т. е. при $S(f) = 0$, (2) превращается в уравнение переноса из § 1 гл. II.

3. Распределение Максвелла и H-теорема. Один из наиболее широко употребляемых интегралов столкновений записывается следующим образом:

$$S(f) = \iiint (f' f_1' - f f_1)\, g b\, db\, d\psi\, d\vec{v}_1. \tag{3}$$

Не приводя громоздкого вывода формулы (3), поясним смысл входящих в нее величин. Через $f = f(\vec{r},\vec{v},t)$, $f_1 = f_1(\vec{r},\vec{v}_1,t)$ обозначены функции распределения до столкновения, через $f' = f'(\vec{r},\vec{v}',t)$, $f_1' = f_1'(\vec{r},\vec{v}_1',t)$ — после столкновения, где $\vec{v}$ и $\vec{v}_1$ — скорости двух сталкивающихся частиц. Величина g — модуль их относительной скорости, b — так называемый прицельный параметр (расстояние наименьшего сближения между частицами), ψ — угловая характеристика их взаимодействия. Интегрирование производится по всем возможным значениям величин b, ψ, $\vec{v}_1$.

Интеграл столкновений в форме Больцмана (3) получается суммированием элементарных актов механического взаимодействия между частицами. При этом предполагается следующее: столкновения упругие (сохраняются суммарные масса, импульс, момент импульса и энергия частиц, что означает, в частности, выполнение равенств $g = g'$, $b = b'$); сила взаимодействия частиц зависит лишь от расстояния

между ними и значительно больше внешних сил; число столкновений с участием более двух частиц пренебрежимо мало (газ не слишком плотный). При известных конкретных законах взаимодействия частиц из (3) следуют более конкретные выражения для интеграла столкновений, используемые, например, при изучении явлений в плазме, где электроны и ионы притягиваются и отталкиваются по закону Кулона.

Используя свойства интеграла (3), получим одно простое, но очень важное решение уравнения Больцмана. Рассмотрим газ, находящийся в состоянии *термодинамического равновесия*, т. е. в ситуации, когда все его макроскопические характеристики постоянны по пространству и не зависят от времени. При этом, очевидно, функция $f(\vec{r},\vec{v},t)$ также не зависит от $\vec{r}$ и t, т. е. $f = f(\vec{v})$. Из (2) сразу же следует уравнение

$$\iiint (f' f_1' - f f_1)\, gb\, db\, d\psi\, d\vec{v}_1 = 0,$$

удовлетворяющееся лишь при выполнении условия

$$f f_1 = f' f_1',$$

или

$$\ln f(\vec{v}) + \ln f_1(\vec{v}_1) = \ln f'(\vec{v}') + \ln f_1'(\vec{v}_1').$$

Другими словами, сумма логарифмов функций распределения сохраняется, являясь инвариантом столкновения. Но при упругих столкновениях инвариантны также полная энергия частиц:

$$\frac{mv^2}{2} + \frac{mv_1^2}{2} = \frac{mv'^2}{2} + \frac{mv_1'^2}{2},$$

и их полный импульс:

$$m\vec{v} + m\vec{v}_1 = m\vec{v}' + m\vec{v}_1'.$$

Эти свойства дают способ нахождения функциональной зависимости f от $\vec{v}$. Ради простоты установим эту зависимость для «одномерного» случая, когда векторы всех скоростей параллельны друг другу. Тогда, умножая последние два равенства на постоянные величины a_1, a_2 и вычитая их из предыдущего, получим

$$\mathcal{F}(v) + \mathcal{F}(v_1) = \mathcal{F}(v') + \mathcal{F}(v_1'),$$

где использовано обозначение $\mathcal{F}(x) = \ln f(x) - a_1 m/2 \cdot x^2 - a_2 m x$. В этом равенстве аргументы слагаемых произвольны, поэтому оно может выполняться лишь при $\mathcal{F}(x) \equiv a_3$ (a_3 — постоянная). Таким образом, для функции распределения справедливо

$$\ln f(v) = \frac{a_1 m}{2} v^2 + a_2 m v + a_3.$$

Выделяя в правой части этого выражения полный квадрат,

$$\ln f(v) = \frac{a_1 m}{2}\left(v + \frac{a_2}{a_1}\right)^2 - \frac{a_2^2 m}{2a_1} + a_3,$$

и избавляясь от логарифма, приходим к формуле

$$f(v) = \exp\left\{\frac{2a_1 a_3 - a_2^2 m}{2a_1}\right\} \exp\left\{\frac{a_1 m}{2}\left(v + \frac{a_2}{a_1}\right)^2\right\}, \tag{4}$$

выражающей зависимость f от v. В силу свойств функции $f(v)$ имеют место (см. (1)) равенства

$$n = \int f\, dv, \quad nV = \int v f\, dv, \quad \frac{3}{2} nkT = \int \frac{m(v-V)^2}{2} f\, dv,$$

где n — концентрация частиц; V — их средняя (макроскопическая) скорость; T — средняя температура частиц (по определению), выражаемая через среднюю кинетическую энергию их хаотического движения с тепловой скоростью $c = v - V$ и измеряемая в градусах Кельвина (см. также п. 1 § 2 гл. II); k — постоянная Больцмана. Используя эти равенства и обозначения, исключим из (4) величины a_1, a_2, a_3:

$$f(v) = n\left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left\{-\frac{m}{2kT}(v-V)^2\right\}. \tag{5}$$

Аналогичное выражение справедливо (при замене v на $\vec{v}$ и V на $\vec{V}$) и в «многомерном» случае. Формула (5) представляет собой одно из решений уравнения Больцмана, описывающее распределение частиц по скоростям в газе, находящемся в состоянии термодинамического равновесия (*распределение Максвелла*). Такое же распределение справедливо не только при полном, но и при *локальном термодинамическом равновесии* (ЛТР), когда макроскопические характеристики n, T, V являются медленно меняющимися функциями $\vec{r}$ и t. Точнее, при ЛТР функции $n(\vec{r},t)$, $T(\vec{r},t)$, $V(\vec{r},t)$ мало меняются на расстояниях порядка длины свободного пробега l и за время свободного пробега τ. Поэтому в условиях ЛТР рассуждения, предшествовавшие выводу формулы (5), применимы к любой точке $\vec{r}$ в любой момент времени t (при этом фигурирующие в (5) значения n, T, V разные в разных точках пространства и в различные моменты времени).

С помощью распределения Максвелла устанавливается имеющая фундаментальное значение *H-теорема Больцмана*, согласно которой с увеличением времени энтропия газа (описываемого уравнением (2)), возрастает. Будем считать распределение частиц по пространству однородным (но зависящим от времени). Тогда уравнение Больцмана (2) принимает вид

$$\frac{\partial f}{\partial t} = \int (f' f_1' - f f_1)\, gb\, db\, d\psi\, d\vec{v}_1. \tag{6}$$

H-функция Больцмана вводится по формуле

$$H(t) = \int f \ln f\, d\vec{v}$$

и по определению представляет собой энтропию единицы объема газа, взятую с обратным знаком, т. е. $S(t) = -H(t)$. Ее производная по времени есть

$$\frac{dH}{dt} = \int (1 + \ln f)\, \frac{\partial f}{\partial t}\, d\vec{v},$$

или, с учетом уравнения (6),

$$\frac{dH}{dt} = \int (1 + \ln f)\,(f' f_1' - f f_1)\, gb\, db\, d\psi\, d\vec{v}\, d\vec{v}_1.$$

Аналогично, принимая во внимание симметрию уравнения (6) относительно функций f и f_1, получаем

$$\frac{dH}{dt} = \int (1 + \ln f_1)\,(f' f_1' - f f_1)\, gb\, db\, d\psi\, d\vec{v}\, d\vec{v}_1,$$

т. е.

$$\frac{dH}{dt} = \frac{1}{2} \int (f' f_1' - f f_1)\,(2 + \ln f f_1)\, gb\, db\, d\psi\, d\vec{v}\, d\vec{v}_1.$$

Вторично используя симметрию уравнения (6) (отражающую обратимость процесса упругих столкновений), нетрудно убедиться в том, что для «обратного» столкновения справедлива такая же формула, т. е.

$$\frac{dH}{dt} = \frac{1}{2} \int (f f_1 - f' f_1')(2 + \ln f' f_1')\, g' b'\, db'\, d\psi'\, d\vec{v}'\, d\vec{v}_1',$$

но в ней величины f', f_1' и f, f_1 поменялись местами. Выше отмечалось, что для упругих столкновений $g = g'$, $b = b'$ (равны также и дифференциалы: $db\, d\psi = db'\, d\psi'$, $d\vec{v}\, d\vec{v}_1 = d\vec{v}'\, d\vec{v}_1'$). Поэтому из двух последних уравнений окончательно получаем

$$\frac{dH}{dt} = -\frac{1}{4} \int (f' f_1' - f f_1) \ln \frac{f' f_1'}{f f_1}\, gb\, db\, d\psi\, d\vec{v}\, d\vec{v}_1. \qquad (7)$$

Знак производной dH/dt определяется знаком функции $\mathcal{F}(x, y) = (x - y)\ln(x/y)$, стоящей под интегралом. Легко устанавливается, что $\mathcal{F}(x, y) \geqslant 0$ (равенство $\mathcal{F}(x, y) = 0$ отвечает равновесному состоянию, когда функция распределения после столкновения не меняется). Поэтому $-dH/dt = dS/dt \geqslant 0$, т. е. энтропия газа в неравновесном состоянии растет, приближаясь к максимальному значению. Оно достигается в состоянии термодинамического равновесия, в котором частицы распределены по закону Максвелла. Таким образом, кинетическое уравнение Больцмана, в отличие от уравнений классической механики, описывает необратимые процессы. Это — следствие перехода от моделей, основанных на первопринципах, к моделям, использующим осредненное статистическое описание совокупности частиц с помощью функции распределения.

4. Уравнения для моментов функции распределения. Следующий после уравнения Больцмана уровень в иерархии математического описания совокупности частиц — гидродинамические модели

газа. Для их получения установим сначала некоторые свойства интеграла столкновений, справедливые не только для больцмановского газа, но и в общем случае. Требуется лишь, чтобы столкновения были упругими. Тогда из неизменности числа частиц до и после столкновения следует равенство

$$\int S(f(\vec{v}))\, d\vec{v} = 0, \qquad (8)$$

а из законов сохранения импульса и энергии вытекают соотношения (упр. 3)

$$\int m\vec{v} S(f(\vec{v}))\, d\vec{v} = 0, \quad \int \frac{mv^2}{2} S(f(\vec{v}))\, d\vec{v} = 0. \qquad (9)$$

В дальнейшем также понадобятся формулы

$$\begin{gathered} \int \Phi(\vec{v}) \frac{\partial f}{\partial t}\, d\vec{v} = \frac{\partial}{\partial t} \int \Phi(\vec{v})\, f\, d\vec{v} = \frac{\partial}{\partial t}(n\langle\Phi\rangle), \\ \int \Phi(\vec{v})\, v_i \frac{\partial f}{\partial x_i}\, d\vec{v} = \frac{\partial}{\partial x_i} \int \Phi(\vec{v})\, v_i f\, d\vec{v} = \frac{\partial}{\partial x_i}(n\langle\Phi(\vec{v})\, v_i\rangle), \end{gathered} \qquad (10)$$

легко выводимые из формулы (1) для средних функций, описывающих состояние газа.

Получим теперь из уравнения Больцмана уравнения для моментов функции распределения, т. е. для величин вида $\int \Phi(\vec{v})\, f\, d\vec{v}$. Ради простоты выкладки будем проводить для одномерного случая, когда $\vec{r} = x$, $\vec{v} = v$. Умножим (2) на $\Phi(\vec{v}) = 1$ и проинтегрируем полученное выражение по скоростям:

$$\int \frac{\partial f}{\partial t}\, dv + \int \frac{\partial f}{\partial x}\, dv + \int S(f)\, dv = 0.$$

Учитывая, что в силу (8) третий член слева в последнем уравнении равен нулю, и используя формулы (10), перепишем последнее уравнение в виде

$$\frac{\partial}{\partial t} n + \frac{\partial}{\partial x}(n\langle v\rangle).$$

Поскольку $\langle v\rangle = V$, то имеем

$$\frac{\partial n}{\partial t} + \frac{\partial}{\partial x}(nV) = 0,$$

или, умножив данное уравнение на m, приходим к уравнению неразрывности

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho V) = 0, \qquad (11)$$

где $\rho = nm$ — плотность газа, V — его макроскопическая скорость. В многомерном случае аналогом (11) является уравнение

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \vec{V}) = 0$$

либо эквивалентное ему уравнение

$$\frac{d\rho}{dt} = -\rho \operatorname{div} \vec{V}. \tag{12}$$

Пусть теперь функция $\Phi(v)$ равна импульсу, т. е. $\Phi(v) = mv$. Повторяя предыдущую процедуру, получим из (2)

$$\int \frac{\partial f}{\partial t} mv\, dv + \int v \frac{\partial f}{\partial x} mv\, dv + \int S(f)\, mv\, dv = 0.$$

В силу первой формулы (9) третий член в левой части этого уравнения равен нулю. Используя (1) и формулы (10), приходим к уравнению

$$\frac{\partial}{\partial t}(nmV) + \frac{\partial}{\partial x}\left(nm\langle v^2\rangle\right) = 0, \tag{13}$$

где $\langle v^2\rangle$ — средний квадрат скорости частиц. Для его вычисления введем хаотическую скорость частиц c (тепловую скорость), представляющую собой разность между собственной скоростью v и скоростью макроскопического движения газа V как единого целого, т. е. $c = v - V$. Для среднего значения v^2 имеем

$$\langle v^2\rangle = \langle (V+c)^2\rangle = V^2 + 2\langle cV\rangle + \langle c^2\rangle = V^2 + \langle c^2\rangle.$$

В последнем равенстве учтено, что $\langle cV\rangle = 0 = V\langle c\rangle$ в силу хаотичности теплового движения. Используя полученное выражение для $\langle v^2\rangle$, из (13) последовательно имеем

$$\frac{\partial}{\partial t}(\rho V) + \frac{\partial}{\partial x}(\rho V^2) + \frac{\partial}{\partial t}\left(\rho\langle c^2\rangle\right) = 0,$$

$$\frac{\partial \rho}{\partial t} + \rho\frac{\partial V}{\partial t} + V\frac{\partial}{\partial x}(\rho V) + \rho V\frac{\partial V}{\partial x} + \frac{\partial}{\partial x}\left(\rho\langle c^2\rangle\right) = 0.$$

Первый и третий члены слева в последнем уравнении в силу (11) взаимно уничтожаются. Поэтому его можно записать как

$$\frac{\partial V}{\partial t} + V\frac{\partial V}{\partial x} = \frac{1}{\rho}\frac{\partial P}{\partial x}. \tag{14}$$

В многомерном случае аналог одномерного уравнения (14) имеет вид

$$\frac{\partial \vec{V}}{\partial t} + \vec{V}(\nabla\vec{V}) = \frac{1}{\rho}\operatorname{grad} P,$$

или, что эквивалентно,

$$\frac{d\vec{V}}{dt} = \frac{1}{\rho}\operatorname{grad} P. \tag{15}$$

В уравнениях движения газа (14), (15) использованы следующие обозначения: $P_{ij} = -\rho\langle c_i c_j\rangle$, $i, j = 1, 2, 3$; $x_i = x, y, z$ для $i = 1, 2, 3$.

В завершение получим уравнение для изменения энергии, взяв в качестве $\Phi(\vec{v})$ величину $mv^2/2$, умножив ее на уравнение (2) и проинтегрировав его по скоростям:

$$\int \frac{\partial f}{\partial t}\frac{mv^2}{2}\, dv + \int v\frac{\partial f}{\partial x}\frac{mv^2}{2}\, dv + \int S(f)\frac{mv^2}{2}\, dv = 0.$$

Третий член слева в данном выражении равен нулю (см. вторую формулу (9)), а первые два записываются по формуле (1) через средние величины:

$$\frac{\partial}{\partial t}\left(n\left\langle\frac{mv^2}{2}\right\rangle\right) + \frac{\partial}{\partial x}\left(n\left\langle\frac{mv^2}{2}v\right\rangle\right) = 0.$$

Снова введем тепловую скорость $c = v - V$ и вычислим величину $\langle v^3\rangle = \langle (V+c)^3\rangle = \langle V^3 + 3V^2c + 3Vc^2 + c^3\rangle = V^3 + 3V\langle c^2\rangle + \langle c^3\rangle$ (здесь учтено, что $\langle 3V^2c\rangle = 3V^2\langle c\rangle = 0$ в силу хаотичности теплового движения, см. также упр. 4). Используя выражения для $\langle v^2\rangle$ и $\langle v^3\rangle$ и равенство $\rho = nm$, получим

$$\frac{1}{2}\frac{\partial}{\partial t}(\rho V^2) + \frac{1}{2}\frac{\partial}{\partial t}\left(\rho\langle c^2\rangle\right) + \frac{1}{2}\frac{\partial}{\partial x}(\rho V^3) + \\ + \frac{3}{2}\frac{\partial}{\partial x}\left(\rho V\langle c^2\rangle\right) + \frac{1}{2}\frac{\partial}{\partial x}\left(\rho\langle c^3\rangle\right) = 0,$$

или, продифференцировав стоящие слева первый и третий члены этого уравнения:

$$\frac{1}{2}\frac{\partial}{\partial t}(\rho V^2) = \frac{V^2}{2}\frac{\partial \rho}{\partial t} + \rho V\frac{\partial V}{\partial t}, \quad \frac{1}{2}\frac{\partial}{\partial x}(\rho V^3) = \frac{V^2}{2}\frac{\partial \rho V}{\partial x} + \rho V^2\frac{\partial V}{\partial x},$$

и принимая во внимание уравнения (11) и (14), умноженные соответственно на V^2 и ρV, придем к уравнению

$$-V\frac{\partial}{\partial x}\left(\rho\langle c^2\rangle\right) + \frac{1}{2}\frac{\partial}{\partial t}\left(\rho\langle c^2\rangle\right) + \frac{3}{2}\frac{\partial}{\partial x}\left(\rho V\langle c^2\rangle\right) + \frac{1}{2}\frac{\partial}{\partial x}\left(\rho\langle c^3\rangle\right) = 0.$$

Представляя в нем третий член слева в виде

$$\frac{3}{2}\frac{\partial}{\partial x}\left(\rho V\langle c^2\rangle\right) = \frac{3}{2}V\frac{\partial}{\partial x}\left(\rho\langle c^2\rangle\right) + \frac{3}{2}\rho\langle c^2\rangle\frac{\partial V}{\partial x},$$

получим

$$\frac{1}{2}V\frac{\partial}{\partial x}\left(\rho\langle c^2\rangle\right) + \frac{1}{2}\frac{\partial}{\partial t}\left(\rho\langle c^2\rangle\right) + \frac{3}{2}\rho\langle c^2\rangle\frac{\partial V}{\partial x} + \frac{1}{2}\frac{\partial}{\partial x}\left(\rho\langle c^3\rangle\right) = 0. \tag{16}$$

Фигурирующая в (16) величина $\rho\langle c^2\rangle/2 = nm\langle c^2\rangle/2$ представляет собой энергию хаотического движения частиц, заключенных в единице объема газа, или его внутреннюю энергию (для простоты считается, что она определяется лишь поступательным движением частиц). Очевидно, что внутренняя энергия ε, отнесенная к единице массы, равна

$\varepsilon = \langle c^2\rangle/2$. Тогда первый и второй члены слева в (16) можно представить как

$$\frac{1}{2}V\frac{\partial}{\partial x}(\rho\langle c^2\rangle) = \rho V\frac{\partial\varepsilon}{\partial x} + \frac{1}{2}V\langle c^2\rangle\frac{\partial\rho}{\partial x},$$

$$\frac{1}{2}\frac{\partial}{\partial t}(\rho\langle c^2\rangle) = \rho\frac{\partial\varepsilon}{\partial t} + \frac{\langle c^2\rangle}{2}\frac{\partial\rho}{\partial t} = \rho\frac{\partial\varepsilon}{\partial t} - \frac{\langle c^2\rangle}{2}\frac{\partial}{\partial x}(\rho V) =$$
$$= \rho\frac{\partial\varepsilon}{\partial t} - \frac{1}{2}\rho\langle c^2\rangle\frac{\partial V}{\partial x} - \frac{1}{2}V\langle c^2\rangle\frac{\partial\rho}{\partial x},$$

где при получении второго выражения использовано уравнение неразрывности (11). Подставив эти равенства в (16), получим

$$\rho\frac{\partial\varepsilon}{\partial t} + \rho V\frac{\partial\varepsilon}{\partial x} + \rho\langle c^2\rangle\frac{\partial V}{\partial x} + \frac{1}{2}\frac{\partial}{\partial x}(\rho\langle c^3\rangle) = 0,$$

или, окончательно,

$$\rho\frac{d\varepsilon}{dt} = P\frac{\partial V}{\partial x} - \frac{\partial W}{\partial x}. \tag{17}$$

В многомерном случае уравнение энергии имеет вид

$$\rho\frac{d\varepsilon}{dt} = -p\operatorname{div}\vec{V} + \sum_{i,j}\Pi_{ij}\frac{\partial V_i}{\partial x_j} - \operatorname{div}\vec{W}. \tag{18}$$

В (17) и (18) величина $\vec{W}$ определяется как

$$\vec{W} = \frac{\rho\langle c^2\vec{c}\rangle}{2} \tag{19}$$

и называется *вектором потока тепла* (в дальнейшем станет понятна оправданность этого названия).

Более сложную структуру имеет введенный выше (см. уравнение (15)) *тензор напряжений* P:

$$P_{ij} = -\rho\langle c_i c_j\rangle = -p\delta_{ij} + \Pi_{ij},$$

где δ_{ij} — символ Кронекера, а величины p, Π_{ij} даются формулами

$$p = \rho\frac{\langle c^2\rangle}{3}, \quad \Pi_{ij} = \rho\left(\frac{\langle c^2\rangle}{3}\delta_{ij} - \langle c_i c_j\rangle\right). \tag{20}$$

Здесь, как нетрудно видеть, p — давление частиц газа, связанное со средней энергией их хаотического движения, т. е. с температурой $T = [m/(3k)]\langle c^2\rangle$, соотношением $p = nkT = \rho kT/m$ (строго говоря, так температура определяется для состояний, близких к равновесным). Второе слагаемое в формуле для P — *тензор вязких напряжений* Π_{ij}. Вязкие силы обусловлены «трением» между частями газа, имеющими разные макроскопические скорости. Его происхождение нетрудно понять, представив себе, например, два плоских слоя газа, движущихся друг относительно друга. Частицы, переходящие через границу раздела из «быстрого» слоя в «медленный» слой, будут увеличивать

среднюю скорость последнего, и наоборот. Тем самым скорости слоев выравниваются, что означает наличие «силы трения» между ними.

Уравнения (11), (12), (14), (15), (17), (18) для моментов записаны относительно средних (гидродинамических) величин, характеризующих газ: плотности, скорости, давления, внутренней энергии. При их выводе не привлекались какие-либо существенные дополнительные предположения, и в этом смысле они принадлежат к тому же иерархическому уровню, что и уравнение Больцмана. Однако, в отличие от уравнения (2), их еще нельзя назвать моделью газа, так как они незамкнуты: кроме ρ, компонент скорости $\vec{V}$, p, ε в них содержатся также величины Π_{ij}, $\vec{W}$, т. е. число уравнений меньше числа неизвестных.

Попытка составить по аналогичной схеме уравнения для моментов более высокого порядка с целью нахождения Π_{ij}, $\vec{W}$ приводит к появлению в этих уравнениях новых неизвестных величин. Поэтому для получения гидродинамических моделей газа необходимо выразить Π_{ij}, $\vec{W}$ через искомые гидродинамические параметры ρ, $\vec{V}$, p, ε. Тогда в пяти уравнениях будут фигурировать пять неизвестных величин (напомним, что внутренняя энергия считается известной функцией плотности и давления: $\varepsilon = \varepsilon(\rho, p)$, или плотности и температуры: $\varepsilon = \varepsilon(\rho, T)$).

Получить эти выражения для Π_{ij}, $\vec{W}$ можно разными способами, например, задавая некоторые полуэмпирические зависимости. Более строгим является способ, состоящий в приближенном «решении» исходного кинетического уравнения (2). Для этого вводится существенное предположение о том, что длина свободного пробега l (время свободного пробега τ) частиц много меньше характерного размера L (характерного времени t) задачи. Другими словами, система близка к состоянию ЛТР, а функция распределения f несильно отличается от локально максвелловской. Тогда, записывая решение в виде бесконечного ряда $f = f^{(0)} + f^{(1)} + \ldots$ (где $f^{(0)}$ — максвелловская функция (5), а $f^{(1)}$ — малое отклонение от нее), считая остальные члены ряда пренебрежительно малыми по отношению к $f^{(1)}$ и подставляя выражение $f = f^{(0)} + f^{(1)}$ в (2), можно найти явный вид возмущения $f^{(1)}$. После этого уже через известную функцию $f = f^{(0)} + f^{(1)}$ по формуле (1) находится тепловая скорость $\vec{c} = \vec{v} + \vec{V}$ и по формулам (19), (20) определяются величины Π_{ij}, $\vec{W}$. Например, $\vec{W}$ вычисляется следующим образом:

$$\vec{W} = \int\frac{m|\vec{c}|^2}{2}\vec{c}(f^{(0)} + f^{(1)})\,d\vec{c}.$$

Заметим, что при $f = f^{(0)}$ вычисления дают $\Pi_{ij} \equiv 0$, $\vec{W} \equiv 0$: в газе с максвелловским распределением частиц вязкие напряжения и потоки тепла отсутствуют.

Не будем описывать сложный и громоздкий процесс нахождения функции $f^{(1)}$ путем «решения» (2) (при его проведении используется

предположение $l \ll L$ $(\tau \ll t)$). Приведем конечный результат:

$$f^{(1)} = \\ = f^{(0)} \left[-A(\xi)\,\vec{c}\sqrt{\frac{m}{2kT}}\frac{\partial \ln T}{\partial \vec{r}} - \frac{m}{2kT}B(\xi)\sum_{i,j}\left(c_i c_j - \frac{1}{3}\delta_{ij}\,|\vec{c}|^2\right)\frac{\partial V_i}{\partial x_j}\right],$$

где $A(\xi)$ и $B(\xi)$ — скалярные функции аргумента $\xi = m\,|\vec{c}|^2/(2kT)$, вид которых зависит, в частности, от вида интеграла столкновений между частицами.

Выражение для $f^{(1)}$ содержит неизвестные макроскопические величины $T(\vec{r},t)$, $V(\vec{r},t)$. Поэтому процедуру определения $f^{(1)}$ нельзя назвать решением уравнения Больцмана в полном смысле этого понятия, т. е. нахождением неизвестной функции $f(\vec{r},\vec{v},t)$ во всей области ее определения. Под решением здесь следует понимать установление с помощью (2) некоторых связей между гидродинамическими параметрами среды.

Использование двух последних формул дает для потока тепла

$$\vec{W} = -\varkappa \operatorname{grad} T. \tag{21}$$

Для компонент тензора вязких напряжений справедливы следующие выражения:

$$\Pi_{ii} = \lambda \operatorname{div}\vec{V} + 2\mu\frac{\partial V_i}{\partial x_i}; \quad \Pi_{ij} = \mu\left(\frac{\partial V_i}{\partial x_j} + \frac{\partial V_j}{\partial x_i}\right), \quad i \neq j. \tag{22}$$

Величина $\varkappa > 0$ называется *коэффициентом теплопроводности*, величины $\lambda > 0$, $\mu > 0$ — *коэффициентами вязкости*. *Коэффициенты переноса* $\varkappa$, λ, μ могут, вообще говоря, зависеть от ρ и T, т. е. от состояния газа. Их функциональная зависимость от ρ, T и численное значение определяются, как следует из вида $f^{(0)}$, $f^{(1)}$, свойствами частиц (массой, характером взаимодействия между ними и т. д.).

Формула (21) — запись закона Фурье для теплопроводных сред (см. также § 2 гл. II), формула (22) — *закон Навье—Стокса* для вязких жидкостей и газов, связывающий компоненты тензора вязких напряжений со *скоростями деформации* (обобщение закона вязкости Ньютона).

5. Цепочка гидродинамических моделей газа. Законы (21), (22) позволяют при заданных коэффициентах переноса замкнуть выведенные в п. 4 уравнения для моментов и получить модели газа в гидродинамическом приближении. Подстановка (21), (22) в (15), (18) приводит (с учетом (12)) к системе пяти уравнений, состоящей из уравнения неразрывности

$$\frac{d\rho}{dt} = -\rho \operatorname{div}\vec{V}, \tag{23}$$

трех уравнений движения (записанных в векторной форме; коэффициенты λ, μ считаются постоянными)

$$\frac{d\vec{V}}{dt} = -\frac{1}{\rho}\operatorname{grad} p + \frac{\lambda+\mu}{\rho}\operatorname{grad}\operatorname{div}\vec{V} + \nu\Delta\vec{V} \tag{24}$$

и уравнения энергии

$$\rho\frac{d\varepsilon}{dt} = -p\operatorname{div}\vec{V} + \lambda(\operatorname{div}\vec{V})^2 + 2\mu\sum_i\left(\frac{\partial V_i}{\partial x_i}\right)^2 + \\ + \mu\sum_{i\neq j}\left(\frac{\partial V_i}{\partial x_j} + \frac{\partial V_j}{\partial x_j}\right)^2 + \operatorname{div}\varkappa\operatorname{grad} T, \tag{25}$$

где $\nu = \mu/\rho$, $df/dt = \partial f/\partial t + (\vec{V}\operatorname{grad})\,f$ — полная (субстанциональная) производная по времени.

При известных краевых условиях из этих уравнений находятся пять искомых функций: ρ, p и три компоненты скорости $\vec{V}$. Система (23)–(25) представляет собой следующий после уравнения Больцмана уровень в иерархии математических описаний большого числа взаимодействующих частиц — модель сжимаемого вязкого теплопроводного газа. Вместо нахождения функции f, зависящей от шести «координат» $\vec{r}$, $\vec{v}$ (а также времени t), и определения через нее средних величин задача свелась к поиску непосредственно гидродинамических параметров, являющихся к тому же функциями лишь трех декартовых координат и времени. Заметим, что с помощью схожих построений из уравнения (2) получаются более сложные, чем (23)–(25), гидродинамические модели, учитывающие, например, анизотропность, наличие внешних сил, химических реакций в среде, заряженных частиц и электромагнитных полей и т. д.

Полученная из исходного кинетического уравнения (2) модель (23)–(25) имеет ясную интепретацию в гидродинамических терминах. Уравнение (23) — запись закона сохранения массы жидкой частицы, полностью совпадающая с уравнением неразрывности, выведенным в § 4 гл. II другим способом. Уравнение (24) выражает второй закон Ньютона, примененный к фиксированной жидкой частице: ее ускорение определяется суммой сил давления и вязких напряжений. Наконец, (25) описывает изменение со временем внутренней энергии жидкой частицы за счет работы сил давления (первый член в правой части), вязкого трения (следующие три члена) и теплопередачи (последний член). В отличие от теплопроводности вязкость, как это сразу видно из (25), всегда приводит к увеличению внутренней энергии всего газа (то же самое справедливо и для энтропии среды).

Модели более низких иерархических уровней получаются из (23)–(25) при соответствующих упрощениях и конкретизациях рассматриваемого объекта, определяемых протекающими в нем физическими процессами, его геометрией и т. д. Число образующихся таким образом иерархических цепочек может быть весьма велико. Продемонстрируем некоторые из них.

В отсутствие теплопроводности, т. е. при $\varkappa = 0$, модель (23)–(25) отвечает вязкому сжимаемому нетеплопроводному газу (*уравнениям Навье—Стокса*, широко применяемым для описания многих процессов в естествознании и технологии). Для важного частного случая несжимаемой жидкости, когда $\rho(\vec{r}, t) = \mathrm{const}$, из (23) следует, что $\operatorname{div}\vec{V} = 0$, и модель Навье—Стокса состоит из четырех уравнений

$$\frac{d\vec{V}}{dt} = -\frac{1}{\rho}\operatorname{grad} p + \frac{\mu}{\rho}\Delta\vec{V}, \quad \operatorname{div}\vec{V} = 0, \tag{26}$$

а уравнение энергии (25) заменяется заданной зависимостью $\varepsilon = \varepsilon(\rho, p) = \varepsilon(p)$.

Когда силы давления гораздо больше сил вязких напряжений, можно положить $\mu = 0$, и из (26) следуют *уравнения движения Эйлера для несжимаемой жидкости*, и вместе с остающимся неизменным уравнением неразрывности получаем

$$\frac{d\vec{V}}{dt} = -\operatorname{grad}\frac{p}{\rho}, \quad \operatorname{div}\vec{V} = 0. \tag{27}$$

Если движение, вдобавок, потенциально, т. е. существует скалярная функция $\phi(x, y, z, t)$ такая, что $\vec{V} = \operatorname{grad}\phi$ (при этом, как следует из известной формулы векторного анализа, $\operatorname{rot}\vec{V} = 0$), то система (27) сводится к *уравнению Лапласа* для потенциала ϕ

$$\Delta\phi = 0,$$

причем первое из уравнений (27) удовлетворяется автоматически (см. упр. 5).

В отсутствие вязкости система (23)–(25) описывает течения *сжимаемого теплопроводного газа*, часто встречающиеся во многих явлениях. В этом случае уравнение энергии (одномерная геометрия)

$$\rho\frac{d\varepsilon}{dt} = -p\frac{\partial V}{\partial x} - \frac{\partial W}{\partial x}, \quad W = -\varkappa\frac{\partial T}{\partial x},$$

совпадает, с учетом перехода от эйлеровых координат к лагранжевым, с уравнением (22) из § 4 гл. II, полученным с помощью прямого применения закона сохранения энергии к жидкой частице. Для неподвижной среды ($\vec{V} \equiv 0$) модель представляет собой *уравнение теплопередачи* (5) из § 2 гл. II

$$C\frac{\partial T}{\partial t} = \operatorname{div}\varkappa(\operatorname{grad} T), \quad C = \rho\frac{\partial\varepsilon(T)}{\partial T},$$

которое при постоянных C, $\varkappa$ для стационарного процесса превращается в уравнение Лапласа для температуры T:

$$\Delta T = 0.$$

Рассмотрим теперь среду без диссипативных процессов вязкости и теплопроводности ($\lambda = 0$, $\mu = 0$, $\varkappa = 0$). Тогда из (23)–(25) сразу

получаются *уравнения движения Эйлера для сжимаемой жидкости*

$$\frac{\partial\vec{V}}{\partial t} + (\vec{V}\operatorname{grad})\vec{V} = -\frac{1}{\rho}\operatorname{grad} p,$$

дополненные уравнениями неразрывности и энергии

$$\frac{d\rho}{dt} = -\rho\operatorname{div}\vec{V}, \quad \rho\frac{d\varepsilon}{dt} = -p\operatorname{div}\vec{V},$$

т. е. приходим к модели, совпадающей с системой (4), (10), (14) из § 4 гл. II, а в случае $\rho = \mathrm{const}$ — с системой (27). Одна из иерархических последовательностей, порождаемая данной моделью, может быть образована, например, следующим путем.

Предположим сначала, что процесс одномерный, и перейдем к массовой координате m. Тогда получим систему (18)–(20) из § 4, гл. II

$$\frac{\partial}{\partial t}\left(\frac{1}{\rho}\right) = \frac{\partial V}{\partial m}, \quad \frac{\partial V}{\partial t} = -\frac{\partial p}{\partial m}, \quad \frac{\partial\varepsilon}{\partial t} = -p\frac{\partial V}{\partial m},$$

где через $\partial/\partial t$ обозначена полная производная по времени. Эта система, состоящая из трех уравнений, в случае изэнтропического течения идеального газа сводится к *квазилинейному уравнению второго порядка* (уравнение (30) из § 4 гл. II):

$$\frac{\partial^2\rho}{\partial t^2} = \frac{\partial}{\partial m}\left(a_0\rho^{\gamma+1}\frac{\partial\rho}{\partial m}\right).$$

Из него для движений, представляющих собой малые отклонения от постоянного течения, следуют *уравнение акустики*

$$\frac{\partial^2\rho}{\partial t^2} = c_0\frac{\partial^2\rho}{\partial m^2}$$

(аналогичное *уравнению колебаний струны* из § 2 гл. III; $c_0 = \sqrt{\gamma p_0/\rho_0}$ — скорость звука), либо, если течение — простая волна, *уравнение Хопфа*

$$\frac{\partial\rho}{\partial t} + \sqrt{a_0}\,\rho^{\frac{\gamma+1}{2}}\frac{\partial\rho}{\partial m} = 0,$$

рассмотренные в п. 7 § 4 гл. II. Наконец, из уравнения акустики в случае течения типа простой волны либо из уравнения Хопфа для малых возмущений вытекает полученное в п. 1 § 1 гл. II *линейное уравнение переноса*

$$\frac{\partial\rho}{\partial t} + \left(a_0\rho_0^{\gamma+1}\right)^{1/2}\frac{\partial\rho}{\partial m} = 0,$$

которое можно считать простейшей моделью газа в классе уравнений в частных производных.

В заключение напомним логику построения иерархии математических описаний большого числа взаимодействующих частиц:

1) наиболее сложные модели основаны на первопринципах, т. е. на применении законов классической механики к каждой частице среды;

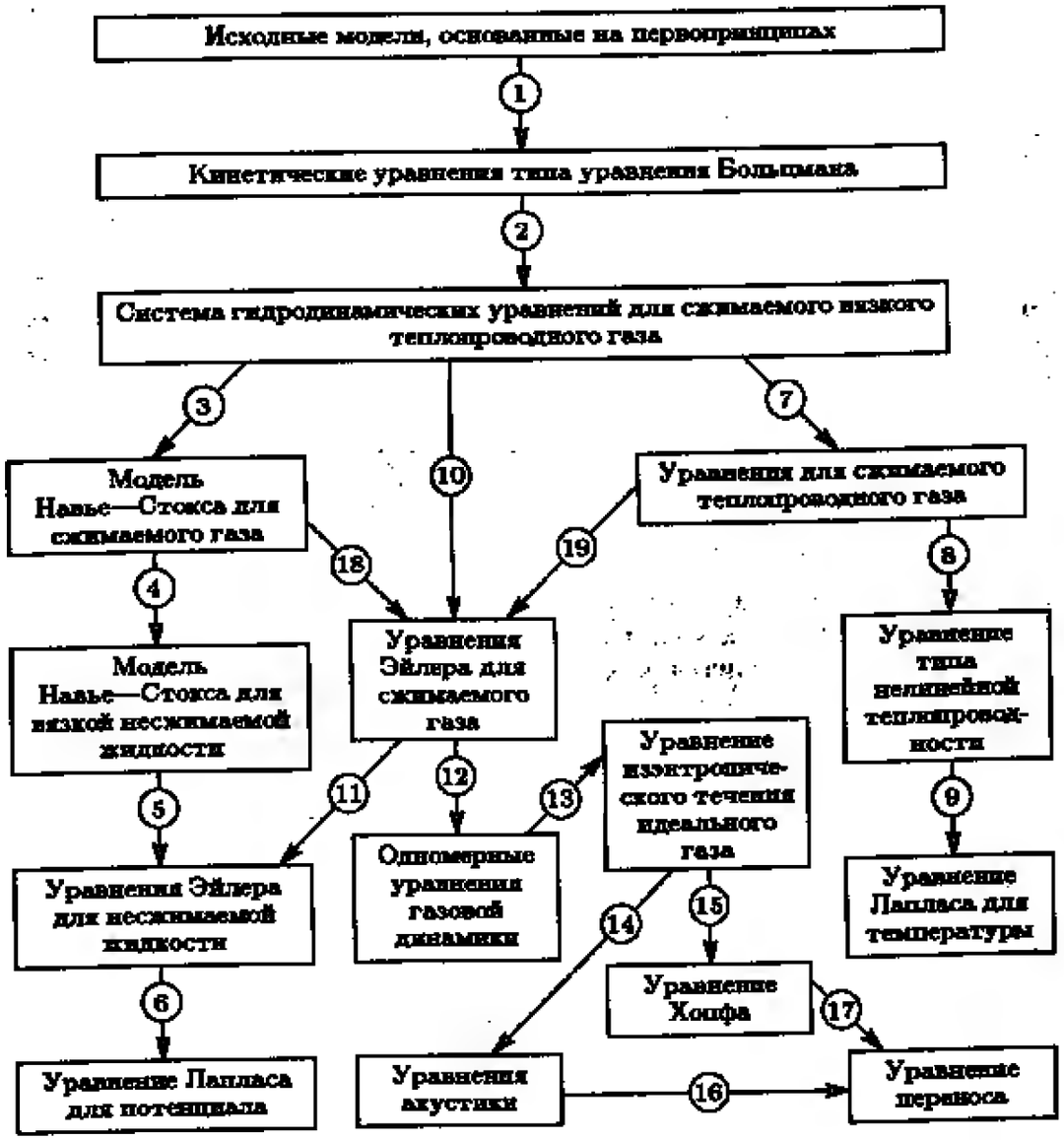


Рис. 41. Иерархические цепочки моделей газа. Стрелки — переход к модели более низкого иерархического уровня, цифры — соответствующие предположения об объекте: 1 — возможность статистического описания газа с помощью функции распределения, упругие столкновения; 2 — близость процессов к локальному термодинамическому равновесию; 3, 19 — отсутствие теплопроводности; 4, 11 — несжимаемость жидкости; 5, 7, 18 — отсутствие вязкости; 6 — идеальная жидкость, потенциальные течения; 8 — неподвижность газа; 9 — стационарный процесс, постоянство коэффициента теплопроводности; 10 — отсутствие вязкости и теплопроводности; 12 — одномерные течения; 13 — идеальный газ, энтропия постоянна; 14, 17 — малые возмущения газа; 15, 16 — течения типа простой волны

2) при переходе к следующему уровню иерархии — кинетическому уравнению Больцмана (2) — используется вероятностное описание газа с помощью функции распределения, а для конкретизации инте-

грала столкновений между частицами вводятся предположения об их упругом характере, об отсутствии тройных столкновений и т. д.;

3) предположение о локальном термодинамическом равновесии — основное для перехода к описанию газа в гидродинамическом приближении, т. е. к уравнениям (23)–(25);

4) основная гидродинамическая модель (23)–(25) порождает, в зависимости от характера рассматриваемых процессов, разнообразные иерархические цепочки, часть из которых была описана выше (рис. 41).

Иерархия моделей газа содержит широкий спектр уравнений, существенно отличающихся друг от друга и в чисто математическом отношении. К ним относятся системы уравнений классической механики высокой размерности, кинетические уравнения, уравнения механики сплошных сред. Последние в свою очередь могут подразделяться на линейные и нелинейные, на уравнения гиперболического типа (уравнения Эйлера и Хопфа), параболического (уравнение теплопроводности) и эллиптического (уравнение Лапласа) типов, а также смешанного типа, на стационарные и нестационарные, многомерные и одномерные уравнения и т. д. Дополнительные вариации построенных моделей связаны с различными вариантами краевых условий и других входных данных.

Продемонстрированный в данном параграфе метод построения моделей по принципу «сверху вниз» более универсален, чем способ, основанный на принципе «снизу вверх». Так, например, уравнение Больцмана невозможно получить из каких-либо моделей более низкого иерархического уровня.

При построении и анализе любой модели всегда полезно знать ее место в общей иерархии моделей изучаемого объекта. Это дает возможность правильно оценивать область ее применимости и четко осознавать ее связи с моделями других уровней, т. е. способствует более глубокому пониманию исследуемых явлений.

УПРАЖНЕНИЯ

1. Используя свойства функции распределения $f(v)$, выражаемые формулой (1), получите из (4) формулу (5).

2. Проверьте, что фигурирующая в (7) функция $\mathcal{F}(x, y)$ неотрицательна.

3. Получите формулы (8), (9), исходя из определения интеграла столкновений $S(f(\vec{v}))$.

4. Убедитесь, используя простые конкретные примеры, в том, что величина $\langle c^3 \rangle$ не равна тождественно нулю (в отличие от средней тепловой скорости $\langle c \rangle$).

5. Используя формулу векторного анализа

$$\frac{1}{2}\,\mathrm{grad}\,|\vec{V}|^2 = \vec{V}\times(\mathrm{rot}\,\vec{V}) + (\vec{V}\nabla)\,\vec{V}$$

и применяя операцию rot к обеим частям первого уравнения (27), получите из него уравнение $\partial(\mathrm{rot}\,\vec{V})/\partial t = \mathrm{rot}[\vec{V}\,\mathrm{rot}\,\vec{V}]$, тождественно удовлетворяющееся в случае потенциальных течений.

Библиография к главе III: [3, 11, 20, 42, 43, 45, 69, 72].

ГЛАВА IV

МОДЕЛИ НЕКОТОРЫХ ТРУДНОФОРМАЛИЗУЕМЫХ ОБЪЕКТОВ

§ 1. Универсальность математических моделей

Построим математические модели динамики скопления амеб и случайного марковского процесса. Поведение «живой» материи (амебы) и «нематериальной» величины (плотность вероятности) опишем теми же параболическими уравнениями, что и явления «мертвой» природы, изученные в гл. II. Рассмотрим также аналогии между некоторыми механическими или физическими объектами и экономическими процессами.

1. Динамика скопления амеб. Амеба — одноклеточный организм размером около десяти микрон (10^{-3} см), обитающий в почве и передвигающийся в ней с помощью ложноножек, т. е. частей своего тела. Питаются амебы в основном бактериями, поглощая их вместе с землей (если пищи достаточно, то амебы размножаются делением на две части).

Из наблюдений и эксперимента известно, что динамика развития их сообщества — достаточно большого количества амеб, находящихся на небольшом расстоянии друг от друга, — бывает достаточно сложной. Например, в зависимости от внешних условий амебы могут собираться в огромные (до сотен тысяч штук) скопления, которые начинают двигаться как единое целое, хотя индивидуальность каждой амебы сохраняется. Замечено, что это макроскопическое «организованное» движение происходит в направлении к более высокой концентрации некоторого химического вещества, вырабатываемого самими амебами. Математическая модель динамики скопления амеб базируется на следующих предположениях:

1) расстояние между амебами мало в сравнении с размерами их скоплений (сотни микрон), их можно рассматривать как «сплошную среду» и вводить концентрацию $N(x,y,z,t)$ — число амеб в единице объема;

2) процесс одномерный, т. е. концентрация амеб и другие величины являются функциями только координаты x и времени t;

3) амебы не рождаются и не умирают в процессе макроскопического движения, т. е. характерное время движения (несколько часов) мало по отношению к характерным временам размножения и жизни амеб;

4) индивидуальное движение амеб при отсутствии стимулирующих внешних воздействий (пища, тепло и т. д.) беспорядочно, хаотично; выделенных направлений нет и каждая амеба может с равной вероятностью двигаться как вправо, так и влево;

5) если в среде есть «притягивающее» химическое вещество, то к собственному неупорядоченному движению амеб добавляется их направленное движение в область с большей плотностью этого вещества.

Составим уравнение баланса амеб в элементе среды dx за время dt, используя «закон сохранения» их числа (предположение 3)). В этом случае общее число амеб в объеме dx (площадь поперечного сечения единична) изменяется лишь из-за разности потока амеб $W(x,t)$ на левой и правой границах элемента. Величина $W(x,t)$ понимается в обычном смысле: это число амеб, пересекающих единичную поверхность за единичное время. Искомое уравнение выглядит так (ср. с уравнением баланса тепла в § 2 гл. II):

$$[\bar{N}(x,t+dt) - \bar{N}(x,t)]\,dx = [\bar{W}(x,t) - \bar{W}(x+dx,t)]\,dt,$$

где $\bar{N}$, $\bar{W}$ — некоторые средние значения величин на малых промежутках dx, dt. Устремляя dx и dt к нулю, приходим к дифференциальному уравнению баланса числа амеб

$$\frac{\partial N}{\partial t} = -\frac{\partial W}{\partial x}.$$

Величина $W = W_c + W_d$ складывается из двух составляющих, W_c и W_d. Часть W_c общего потока формируется за счет хаотического движения амеб, и поэтому по аналогии с законом Фурье для процесса диффузии тепла (§ 2 гл. II) его можно записать через градиент их концентрации:

$$W_c = -\mu\frac{\partial N}{\partial x},$$

где $\mu > 0$ — некоторый коэффициент, характеризующий рассматриваемую «среду». Эту формулу для W_c и величину μ нетрудно получить из более подробного анализа процесса на микроуровне, используя те же рассуждения, которые применялись для явлений передачи тепла.

При получении выражения для составляющей W_d, описывающей направленный поток амеб, будем считать, что величина W_d тем больше, чем больше градиент плотности «притягивающего» вещества:

$$W_d = \eta N\frac{\partial \rho}{\partial x}.$$

Здесь $\eta > 0$ — некоторая постоянная, $\rho(x,t)$ — плотность вещества, а множитель N перед градиентом означает, что при заданном градиенте величины ρ составляющая потока W_d пропорциональна концентрации амеб в данной точке. Объединяя выражения для W_c, W_d и подставляя их в уравнение баланса, получаем

$$\frac{\partial N}{\partial t} = \frac{\partial}{\partial x}\left(\mu\frac{\partial N}{\partial x} - \eta N\frac{\partial \rho}{\partial x}\right). \qquad (1)$$

В уравнении (1) две неизвестных функции — N и ρ. Поэтому необходимо получить, пользуясь законом сохранения вещества, уравнение баланса для величины ρ. При этом следует учесть, что скорость выделения химического вещества пропорциональна концентрации амеб. Будем учитывать также распад вещества, скорость которого, естественно, пропорциональна его концентрации (аналогично процессу радиоактивного распада; см. § 1 гл. I). Таким образом, в единицу времени в единичном объеме появляется и исчезает количество вещества, равное

$$f = \alpha N - \beta\rho,$$

где $\alpha > 0$, $\beta > 0$ — константы, характеризующие соответственно скорость его выделения амебами и скорость распада (в этом состоит отличие модели для вещества от модели для амеб, в которой они не умирают и не рождаются). Изменение плотности вещества в элементарном объеме среды происходит также и вследствие разности его потоков на левой и правой границах элемента. Оно диффундирует в среде из мест с большей концентрацией в места с меньшей концентрацией подобно тому, как тепло распространяется от более нагретых участков теплопроводной среды к менее нагретым. Это движение создает, согласно *закону Фика*, поток W_ρ, равный

$$W_\rho = -D\frac{\partial\rho}{\partial x},$$

где $D > 0$ — коэффициент диффузии (вывод закона Фика аналогичен выводу закона Фурье).

Итак, уравнение баланса вещества имеет вид

$$\frac{\partial\rho}{\partial t} = -\frac{\partial W_\rho}{\partial x} + f,$$

или, учитывая выражения для W_ρ и f,

$$\frac{\partial\rho}{\partial t} = D\frac{\partial^2\rho}{\partial x^2} + \alpha N - \beta\rho. \qquad (2)$$

Уравнения (1), (2) вместе со входными данными μ, η, α, β, D служат моделью динамики скопления амеб при сделанных выше предположениях. Для однозначного определения решения, т. е. функций $N(x,t)$, $\rho(x,t)$, необходимо, как обычно, знать соответствующие начальные и граничные условия. Наиболее просто они выглядят в случае задачи Коши, когда процесс рассматривается в неограниченном пространстве $-\infty < x < \infty$; тогда достаточно задать в момент $t = 0$ начальную концентрацию амеб $N(x,0) = N_0(x)$ и плотность вещества $\rho(x,0) = \rho_0(x)$.

Уравнения (1), (2) взаимосвязаны: в первое из них входит величина ρ, во втором фигурирует величина N. Система (1), (2) нелинейна из-за наличия члена $\eta N\,\partial\rho/\partial x$ в скобках, стоящих в правой части (1). Рассматриваемые соответственно относительно концентраций амеб N и плотности вещества ρ уравнения (1) и (2), как нетрудно видеть, принадлежат к параболическому типу.

Если амебы не выделяют «притягивающего» вещества и $\rho(x,t) \equiv 0$, то (1) переходит в уравнение теплопроводности (или диффузии)

$$\frac{\partial N}{\partial t} = \mu\frac{\partial^2 N}{\partial x^2},$$

что легко объяснить, поскольку в потоке W остается лишь составляющая W_c, соответствующая беспорядочному, ненаправленному движению амеб. Когда амебы по каким-то причинам перестают выделять вещество, коэффициент α в (2) становится равным нулю, и с этого момента уравнение (2) принимает вид (диффузия с распадом)

$$\frac{\partial\rho}{\partial t} = D\frac{\partial^2\rho}{\partial x^2} - \beta\rho \qquad (3)$$

и сводится простой заменой (упр. 1) также к уравнению теплопроводности

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2}.$$

Из-за нелинейности системы (1), (2) сконструировать ее общее решение нельзя, и поэтому определение пространственно-временной динамики скопления амеб — весьма непростая задача. Однако она значительно облегчается, если изучаются малые отклонения от постоянного во времени и по пространству решения $N \equiv N_0$, $\rho \equiv \rho_0$, т. е. когда нелинейная задача становится линейной. Такое решение существует лишь при соотношении

$$\alpha N_0 = \beta\rho_0,$$

означающем, что выделение вещества и его распад уравновешивают друг друга.

Линеаризованная в окрестности постоянного решения система (1), (2) (ср. с уравнением акустических колебаний в § 4 гл. II) имеет вид

$$\begin{aligned} \frac{\partial\tilde N}{\partial t} &= \frac{\partial}{\partial x}\left(\mu\frac{\partial\tilde N}{\partial x} - \eta N_0\frac{\partial\tilde\rho}{\partial x}\right), \\ \frac{\partial\tilde\rho}{\partial t} &= D\frac{\partial^2\tilde\rho}{\partial x^2} + \alpha\tilde N - \beta\tilde\rho, \end{aligned} \qquad (4)$$

где $\tilde N$ и $\tilde\rho$ — малые возмущения ($\tilde N \ll N_0$, $\tilde\rho \ll \rho_0$). Ее общее решение для неограниченного пространства $-\infty < x < \infty$ (в этом случае нет необходимости удовлетворять граничным условиям) можно построить в виде суммы частных решений (гармоник)

$$\tilde N = C_1 \sin kx\, e^{\gamma t}, \quad \tilde\rho = C_2 \sin kx\, e^{\gamma t},$$

где $k > 0$ — волновое число, C_1, C_2 — константы. Для частных решений должны выполняться соотношения

$$\begin{aligned} C_1(\gamma + \mu k^2) &= C_2\eta N_0 k^2, \\ C_2(\gamma + \beta + Dk^2) &= C_1\alpha, \end{aligned} \qquad (5)$$

связывающие длину волны гармоники $\lambda = 2\pi/k$ с величиной γ — ее инкрементом (или декрементом), характеризующим нарастание или затухание возмущения со временем. Исключая C_1 и C_2 из (5), получаем относительно γ квадратное уравнение

$$\gamma^2 + b\gamma + c = 0, \tag{6}$$

где $b = \beta + k^2(\mu + D)$, $c = \mu k^2(\beta + Dk^2) - \eta\alpha N_0 k^2$. Оба корня уравнения (6) отрицательны тогда и только тогда, когда $c > 0$, т. е. при выполнении неравенства

$$\mu(\beta + Dk^2) > \eta\alpha N_0. \tag{7}$$

Если имеет место (7), то для любых значений k амплитуда возмущений с любой длиной волны уменьшается с течением времени, и постоянное решение устойчиво. Неравенство (7) заведомо выполнено при

$$\mu\beta > \eta\alpha N_0 \quad (\text{или } \mu > \eta\rho_0),$$

т. е. для фиксированных параметров задачи при достаточно малых концентрациях амеб (и плотностях «притягивающего» вещества). В противном случае возможна неустойчивость постоянного решения (если в начальном спектре возмущений найдутся длинноволновые гармоники с небольшими значениями k, растущие со временем), порождающая более сложную картину эволюции скопления амеб. Разумеется, линеаризованная модель не дает исчерпывающей картины процесса, но из нее можно извлечь ряд сведений, полезных для более полного исследования.

2. Случайный марковский процесс. Типичным примером подобного процесса служит движение помещенной в жидкость маленькой твердой частицы, совершающей хаотические перемещения под действием беспорядочных столкновений с молекулами жидкости (броуновское движение). Ее положение в любой момент времени $t \geqslant t_0$ задается координатами x, y, z трехмерного пространства $\mathbb{R}^3$. В дальнейшем для упрощения выкладок будем рассматривать одномерное движение, т. е. случайные блуждания броуновской частицы вдоль оси x, $x \in \mathbb{R}^1$.

Случайный процесс называется *марковским*, если по положению точки x в момент времени t однозначно определяется вероятность ее нахождения в произвольный момент $t' > t$ в некоторой части (в любом измеримом подмножестве E) пространства $\mathbb{R}^1$. Другими словами, это процесс без последействия, когда события, случившиеся в промежутке времени между t и t', не влияют на положение точки в момент t'.

Марковский процесс полностью характеризуется функцией

$$p(t, x, t', x'), \qquad x \in \mathbb{R}^1,$$

называемой *плотностью вероятности* в точке x', зная которую, нетрудно вычислить вероятность

$$p(x, t, t', E) = \int\limits_{E(x')} p(x, t, x', t')\, dx'$$

нахождения частицы в некоторой окрестности $E(x')$ точки x' в момент времени t'.

Очевидно, что для функции p выполнено условие нормировки

$$\int\limits_{\mathbb{R}^1} p(t, x, t', x')\, dx' = 1, \tag{8}$$

т. е. в любой момент частица обязательно находится в какой-либо точке пространства $\mathbb{R}^1$.

При построении модели марковского процесса существенным образом используется предположение о его *сильной непрерывности*. Считается, что частица за малые промежутки времени Δt может получить заметные приращения координаты $\Delta x \geqslant \delta$ лишь с малой вероятностью. Это значит, что для любого $\delta > 0$

$$\int\limits_{|x'-x| \geqslant \delta} p(t - \Delta t, x, t, x')\, dx' = o(\Delta t),$$

или, в эквивалентной записи,

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int\limits_{|x'-x| \geqslant \delta} p(t - \Delta t, x, t, x')\, dx' = 0. \tag{9}$$

Предполагается также, что для любых $\delta > 0$ существуют равномерные по x пределы

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int\limits_{|x'-x| < \delta} (x' - x)\, p(t - \Delta t, x, t, x')\, dx' = b > 0, \tag{10}$$

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int\limits_{|x'-x| < \delta} (x' - x)^2\, p(t - \Delta t, x, t, x')\, dx' = 2a > 0. \tag{11}$$

Эти предположения имеют следующую интерпретацию: вероятность для частицы находиться в момент t в интервале $|x' - x| < \delta$ пропорциональна Δt (уменьшается с уменьшением промежутка времени, прошедшего от начального момента $t - \Delta t$, что естественно) и обратно пропорциональна некоторой «средней» величине размера интервала $|x' - x|$ (и его квадрату, что также естественно). Величины a и b зависят, вообще говоря, от точки x и момента t, т. е. $a = a(x, t)$, $b = b(x, t)$, но для простоты здесь рассматривается частный случай, когда a и b — постоянные.

Наконец, последнее используемое ниже предположение состоит в том, что для любых t, x, t', x' существуют непрерывные частные производные функции p по x

$$\frac{\partial p}{\partial x}, \quad \frac{\partial^2 p}{\partial x^2} \tag{12}$$

Основное свойство рассматриваемого процесса выражается *тождеством Маркова*

$$p(t,x,t',x') = \int\limits_{\mathbb{R}^1} p(x,t,\bar{t},\bar{x})\, p(\bar{t},\bar{x},t',x')\, d\bar{x}, \tag{13}$$

где $t < \bar{t} < t'$ — некоторый момент времени в промежутке от t до t', а $\bar{x}$ — координата частицы в момент $\bar{t}$, $t < \bar{t} < t'$, некоторая промежуточная точка при движении от точки x до точки x'. Смысл тождества (13) проясняется при рассмотрении перехода из точки x в точку x' как последовательности двух переходов — сначала из точки x в точку $\bar{x}$, а затем из точки $\bar{x}$ в точку x' (на рис. 42 показаны возможные комбинации этих переходов, причем промежуточные перемещения изображены штриховыми линиями, а основной переход — сплошными). Вероятность события, состоящего из двух последовательных независимых событий, равна произведению вероятностей каждого события, поэтому под интегралом в (13) стоит произведение соответствующих величин. Интеграл берется по всем возможным промежуточным точкам $\bar{x} \in \mathbb{R}^1$.

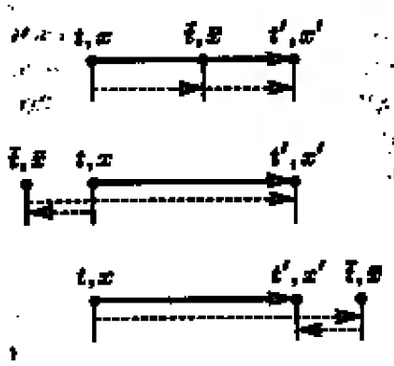


Рис. 42

Для рассматриваемого процесса тождество Маркова играет роль своеобразного «фундаментального закона», связывая некоторым образом значения функции p в точках t,x и t',x'. С его помощью вычислим сначала разность величин p в моменты $t-\Delta t$ и t:

$$p(t-\Delta t,x,t',x') - p(t,x,t',x') =$$

$$= \int\limits_{\mathbb{R}^1} p(t-\Delta t,x,t,\bar{x})\, p(t,\bar{x},t',x')\, d\bar{x} - p(t,x,t',x') \int\limits_{\mathbb{R}^1} p(t-\Delta t,x,t,\bar{x})\, d\bar{x}.$$

Первому члену в левой части этого равенства отвечает (см. (13)) первый интеграл в его правой части, а второму — тождественный ему член, умноженный на интеграл, который согласно условию нормировки (8) равен единице. Множитель $p(t,x,t',x')$ не зависит от $\bar{x}$, и поэтому, внося его под знак интеграла, перепишем равенство в виде

$$p(t-\Delta t,x,t',x') - p(t,x,t',x') =$$

$$= \int\limits_{\mathbb{R}^1} [p(t,\bar{x},t',x') - p(t,x,t',x')]\, p(t-\Delta t,x,t,\bar{x})\, d\bar{x}.$$

Поделим теперь в этом равенстве обе его части на Δt и разобьем интеграл на два — по области $|\bar{x}-x| \geqslant \delta$ и $|\bar{x}-x| < \delta$:

$$\frac{p(t-\Delta t,x,t',x') - p(t,x,t',x')}{\Delta t} = I_1 + I_2, \tag{14}$$

$$I_1 = \frac{1}{\Delta t} \int\limits_{|\bar{x}-x| \geqslant \delta} [p(t,\bar{x},t',x') - p(t,x,t',x')]\, p(t-\Delta t,x,t,\bar{x})\, d\bar{x},$$

$$I_2 = \frac{1}{\Delta t} \int\limits_{|\bar{x}-x| < \delta} [p(t,\bar{x},t',x') - p(t,x,t',x')]\, p(t-\Delta t,x,t,\bar{x})\, d\bar{x}.$$

Интеграл I_1 в силу свойства сильной непрерывности (9) стремится к нулю при $\Delta t \to 0$ (первый сомножитель в подынтегральном выражении не зависит от Δt и не влияет на поведение I_1 при $\Delta t \to 0$).

Интеграл I_2 преобразуем, раскладывая с учетом (12) первый сомножитель по степеням $(\bar{x}-x)$:

$$I_2 = \frac{1}{\Delta t} \int\limits_{|\bar{x}-x| < \delta} \frac{\partial p(t,x,t',x')}{\partial x} (\bar{x}-x)\, p(t-\Delta t,x,t,\bar{x})\, d\bar{x} +$$

$$+ \frac{1}{\Delta t} \int\limits_{|\bar{x}-x| < \delta} \frac{1}{2} \frac{\partial^2 p(t,x,t',x')}{\partial x^2} (\bar{x}-x)^2\, p(t-\Delta t,x,t,\bar{x})\, d\bar{x} +$$

$$+ \frac{1}{\Delta t} \int\limits_{|\bar{x}-x| < \delta} o[(\bar{x}-x)^2]\, \frac{1}{2} \frac{\partial^2 p(t,x,t',x')}{\partial x^2}\, p(t-\Delta t,x,t,\bar{x})\, d\bar{x}.$$

Устремим в этом равенстве Δt к нулю, замечая, что частные производные функции p, стоящие под знаком интеграла, не зависят от $\bar{x}$. Вынося их из-под интеграла, для первых двух членов получим, в силу предположений (10), (11), что их пределы соответственно равны

$$b\,\frac{\partial p(t,x,t',x')}{\partial x}, \quad a\,\frac{\partial^2 p(t,x,t',x')}{\partial x^2},$$

а третье слагаемое представим в виде

$$\bar{\varepsilon}(\bar{x}-x)\, \frac{1}{2} \frac{\partial^2 p(t,x,t',x')}{\partial x^2} \frac{1}{\Delta t} \int\limits_{|\bar{x}-x| < \delta} (\bar{x}-x)^2\, p(t-\Delta t,x,t,\bar{x})\, d\bar{x},$$

где $\bar{\varepsilon}(\bar{x}-x)$ — среднее значение функции $\varepsilon(\bar{x}-x)$, причем по определению величины $o[(\bar{x}-x)^2]$ имеем $\bar{\varepsilon}(\bar{x}-x) \to 0$ при $\delta \to 0$. Осуществим в последнем выражении предельный переход сначала при $\delta \to 0$. Этот предел, очевидно, равен нулю при любых $\Delta t > 0$. Поэтому его предел при $\Delta t \to 0$ также равен нулю. Наконец, предел левой части (14) при $\Delta t \to 0$ равен производной функции p по времени. Суммируя эти результаты, получаем из (14) *уравнение Колмогорова* для плотности вероятности, справедливое для всех $t > t_0$, $-\infty < x < \infty$:

$$\frac{\partial p(t,x,t',x')}{\partial t} = a\,\frac{\partial^2 p(t,x,t',x')}{\partial x^2} + b\,\frac{\partial p(t,x,t',x')}{\partial x}. \tag{15}$$

Уравнение (15) — линейное параболическое уравнение. Этими же свойствами обладают и его обобщения. Например, в случае, когда величины b и a в (10), (11) зависят от t, x, аналог уравнения (15) имеет вид

$$\frac{\partial p}{\partial t} = a(t,x)\frac{\partial^2 p}{\partial x^2} + b(t,x)\frac{\partial p}{\partial x}. \tag{16}$$

Если же точка x принадлежит n-мерному пространству, т. е. $x = (x_1, x_2, \ldots, x_n)$, то для функции p справедливо следующее обобщение (15), (16):

$$\frac{\partial p}{\partial t} = \sum_{i,j=1}^{n} a_{ij}(t,x)\frac{\partial^2 p}{\partial x_i\,\partial x_j} + \sum_{i=1}^{n} b_i(t,x)\frac{\partial p}{\partial x_i}, \tag{17}$$

где $b_i(t,x)$ и $a_{ij}(t,x)$ вычисляются по формулам (10), (11), но в (10) вместо сомножителя $(x'-x)$ берется величина $(x'-x)_i$, а в (11) вместо $(x'-x)^2$ фигурирует выражение $(x'-x)_i\,(x'-x)_j$, $i,j = 1,2,\ldots,n$.

Заметим, что (17) отнюдь не формальное обобщение (15), (16). Случайные марковские процессы могут протекать не только в реальном физическом пространстве (броуновское движение), но и в так называемом фазовом пространстве. Они характерны для многих технических и иных систем, состояние которых описывается совокупностью фазовых переменных $x_1, x_2, \ldots, x_n$, число которых может быть значительно больше трех.

Простейший вариант уравнения Колмогорова получается из (15) при $b = 0$:

$$\frac{\partial p}{\partial t} = a\frac{\partial^2 p}{\partial x^2}, \tag{18}$$

и представляет собой *уравнение теплопроводности* (или диффузии).

Однако между математическими моделями теплопередачи и случайного марковского процесса есть существенная разница. При выводе уравнения (15) для функции $p(t,x)$ использовались (в отличие от вывода уравнения теплопроводности) условия сильной непрерывности (9)–(11). Это приводит к тому, что функция $p(t,x)$ не может быть любым решением уравнения Колмогорова. Она оказывается так называемым *фундаментальным решением уравнений* (15)–(18).

Проще всего пояснить это свойство функции $p(t,x)$ в случае простейшего уравнения (18). Пусть функция $u(t,x)$ — решение уравнения (18), определенное при $t > t_0$, $-\infty < x < \infty$ и удовлетворяющее заданному начальному условию

$$u(t,x) \to u_0(x) \geqslant 0 \quad \text{при} \quad t \to t_0. \tag{19}$$

Тогда, если $p(x,t,t',x')$ — фундаментальное решение (18), то функция $u(t,x)$ находится по формуле

$$u(t,x) = \int\limits_{\mathbb{R}^1} p(t,x,t',x')\,u_0(x')\,dx'. \tag{20}$$

То, что $u(t,x)$, задаваемая в виде (20), — решение (18), нетрудно установить дифференцированием (упр. 4). Свойство (19) доказывается с помощью разбиения интеграла в (20):

$$u(t,x) = \int\limits_{|x'-x|<\delta} p(t,x,t',x')\,u_0(x')\,dx' + \int\limits_{|x'-x|\geqslant\delta} p(t,x,t',x')\,u_0(x')\,dx'.$$

При $t = t_0$ точка имеет координату x, и поэтому по свойству сильной непрерывности вероятность ее нахождения при $t' \to t \to t_0$ в области $|x'-x| \geqslant \delta$ нулевая, т. е. $p(t,x,t',x') \to 0$ при $t' \to t_0$. Следовательно, из последней формулы получаем

$$\lim_{t'\to t_0}\int\limits_{\mathbb{R}^1} p(t,x,t',x')\,u_0(x')\,dx' = \lim_{t'\to t_0}\int\limits_{|x-y|<\delta} p(t,x,t',x')\,u_0(x')\,dx'.$$

Используя условие нормировки (8) при $t' \to t_0$, $x' \to x$ и учитывая независимость левого предела в последней формуле от δ, получаем

$$\lim_{t'\to t_0}\int\limits_{\mathbb{R}^1} p(t,x,t',x')\,u_0(x')\,dx' = u_0(x),$$

т. е. формулу (19).

Тот факт, что функция $p(t,x)$ не любое, а именно фундаментальное решение уравнения Колмогорова (для уравнений (15)–(17) доказательство аналогично), не является дефектом рассматриваемой модели, а отражает естественное свойство случайного марковского процесса. Действительно, в начальный момент $t = t_0$ блуждающая точка имеет некоторую координату x_0, и поэтому $p(x,t_0) \equiv 0$ при $x \neq x_0$. В то же время из условия (8) при $t' = t = t_0$ следует, что

$$\int\limits_{\mathbb{R}^1} p(x,t_0)\,dx = 1,$$

т. е. начальными данными для уравнения (18) (и уравнений (15)–(17)) служит δ-функция. Решения задачи Коши для линейных параболических уравнений с таким начальным условием являются их фундаментальными решениями. Простейший пример — функция мгновенного точечного источника тепла для уравнения теплопроводности (11) из § 2 гл. II, даваемая простой формулой (20) из § 2 гл. II и имеющая те же свойства, что и функция $p(t,x)$. Для более общих уравнений (15)–(17) не существует столь простых представлений их фундаментальных решений. Однако то обстоятельство, что величина $p(t,x)$ подчиняется уравнению Колмогорова и является его фундаментальным решением, существенно используется при исследовании объектов, в которых протекают случайные марковские процессы, в частности в задачах управления такими объектами.

Таблица 1

Универсальность математических моделей. Параболические уравнения

Объект (процесс)	Основные предположения и законы
Движение грунтовых вод	Сохранение массы, закон Дарси
Теплопередача; диффузия вещества	Сохранение энергии, закон Фурье; сохранение массы, закон Фика
Движение скопления амеб	Сохранение числа амеб, хаотичность движения амеб в отсутствие «притягивающего» вещества
Случайный марковский процесс	Тождество Маркова, сильная непрерывность процесса
Динамика распределения власти в иерархии	Законопослушность, постулат о механизмах перераспределения власти в иерархии

Итак, параболические уравнения — один из примеров универсальности математических моделей (табл. 1). Они описывают широкий круг процессов совершенно различной природы (последний пример в таблице рассматривается в § 4). Отметим, что параболические уравнения часто связываются с хаотическими неупорядоченными явлениями (теплопередача, диффузия и т. д.). Однако они применимы и ко многим процессам, которые рассматриваются как детерминированные (движение грунтовых вод, фильтрация газа в пористой среде и т. д.).

Универсальность математических моделей — отражение единства окружающего нас мира и способов его описания. Поэтому методы и результаты, разработанные и накопленные при математическом моделировании одних явлений, относительно легко, «по аналогии», могут быть перенесены на широкие классы совсем других процессов.

3. Примеры аналогий между механическими, термодинамическими и экономическими объектами. Особое значение «метод аналогий» имеет при математическом моделировании трудноформализуемых объектов, для которых фундаментальные законы, вариационные принципы и иные общие и математически строгие утверждения либо неизвестны, либо вообще не существуют. К таким объектам относятся, например, системы с заметным вмешательством людей, в частности экономические системы.

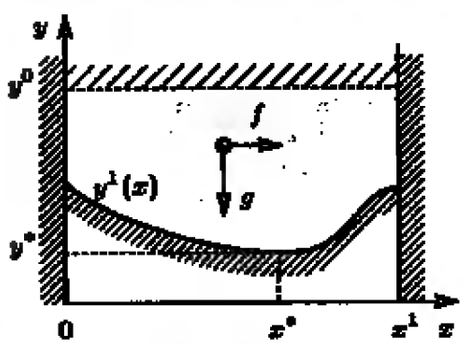


Рис. 43

Одна из важных механико-экономических аналогий — аналогия между равновесием материальной частицы в потенциальном поле внешних сил и выбором оптимального плана производства. Рассмотрим ради простоты ее частный случай. Пусть единичная масса (точка на рис. 43) находится в поле силы тяжести и

может занимать, вообще говоря, любое положение в области пространства, ограниченного снизу твердой идеальной поверхностью. Уравнение поверхности дается зависимостью $y^1(x) \geqslant 0$, где x, y — соответственно горизонтальная и вертикальная координаты частицы, $y^1(x)$ — гладкая функция. Очевидно, что величина y — расстояние частицы от оси абсцисс — подчиняется неравенству

$$y \geqslant y^1(x), \qquad 0 \leqslant x \leqslant x^1. \tag{21}$$

Поле силы тяжести потенциально, т. е. существует функция (потенциал) $P(x, y)$ такая, что компоненты ее градиента определяют внешнюю силу, действующую на материальную частицу в данной точке поля. В рассматриваемой ситуации

$$\operatorname{grad} P = (0, -g),$$

где g — ускорение свободного падения (горизонтальная компонента, естественно, равна нулю). Следовательно, потенциал дается формулой

$$P = -gy, \tag{22}$$

в которой опущена несущественная аддитивная постоянная.

Изменение потенциала при перемещениях частицы в поле силы тяжести равно работе A, совершаемой этой силой, и определяется только начальным и конечным положением частицы (в данном случае только координатой y):

$$A = \int\limits_{x_1, y_1}^{x_2, y_2} dP = P(x_2, y_2) - P(x_1, y_1) = P(y_2) - P(y_1). \tag{23}$$

Точка x^*, y^* называется *положением равновесия*, если частица, помещенная в нее и имеющая нулевую скорость, остается в ней любое время (при наложенных связях (21) и внешней силе $\operatorname{grad} P(x^*, y^*)$). Потенциал в точке равновесия достигает своего экстремального значения. Это свойство хорошо иллюстрирует рис. 43, где равновесными являются точки минимума функции $y^1(x)$ (в них сила тяжести уравновешивается реакцией опоры). Любое *виртуальное* (не противоречащее связям (21)) малое смещение частицы из этих точек приведет к тому, что над нею будет совершена, как видно из (23), отрицательная работа, и потенциал уменьшится (если перемещение полностью или частично осуществляется по поверхностям $y^1(x)$, $x = 0$, $x = x^1$, то, поскольку они считаются идеальными и силы трения отсутствуют, а их реакция перпендикулярна перемещению, это не сказывается на величине работы). Во всех других точках области $y \geqslant y^1(x)$, $0 \leqslant x \leqslant x^1$ потенциал (22) не достигает своего максимального значения.

Итак, поиск устойчивого положения равновесия сводится к решению задачи

$$P(x, y) \to \max \quad \text{при} \quad y \geqslant y^1(x), \quad 0 \leqslant x \leqslant x^1. \tag{24}$$

Закон покоя Мопертюи (24) аналогичным образом формулируется и для общих механических систем, точки равновесия которых находятся не столь просто, как в приведенном случае. Например, если

на частицу действуют, кроме силы тяжести, другие потенциальные силы, то точки равновесия отнюдь не обязаны совпадать с точками минимумов функции $y^1(x)$.

В экономической интерпретации задача (24) называется *задачей нелинейного программирования* и часто возникает при планировании производства.

Пусть некоторое предприятие выпускает продукцию (кирпичи), объем которой обозначим x, $0 \leqslant x \leqslant x^1$. Для производства предприятию необходимо затратить некоторый ресурс (глину), остаток которого после выполнения плана обозначим y (начальный ресурс равен y^0 и считается независящим от величины плана x, $y^0 \leqslant y \leqslant 0$). Известно, что производство происходит при некоторых ресурсных ограничениях сверху, т. е.

$$y^0 \leqslant y \leqslant y^1(x), \quad 0 \leqslant x \leqslant x^1, \tag{25}$$

где $y^1(x)$ — минимальное количество неиспользованных ресурсов, которое предприятие по технологическим, финансовым или иным причинам обязано иметь в своем распоряжении после выполнения плана x ($y^1(x)$ считается заданной гладкой функцией x).

В упрощенной постановке прибыль $P(x, y)$ равна разности между стоимостью выпущенной продукции fx и стоимостью затраченных ресурсов $g(y - y^0)$:

$$P(x, y) = fx + g(y^0 - y), \tag{26}$$

где f, g — цена единицы продукции и ресурса соответственно (остальные затраты считаются несущественными).

Задача планирования производства состоит в том, чтобы выбрать план x^*, выполнение которого дает максимальную прибыль (26) при ресурсных ограничениях (25):

$$P(x, y) \to \max \quad \text{при} \quad y^0 \leqslant y \leqslant y^1(x), \quad 0 \leqslant x \leqslant x^1. \tag{27}$$

Формулировки, подобные (27), справедливы для весьма общих задач планирования.

Для получения полной аналогии несколько усложним задачу о равновесии материальной частицы. Введем дополнительно к силе тяжести потенциальную внешнюю силу (см. рис. 43), действующую на частицу в направлении оси x и равную по величине f. Например, для заряженной частицы эта сила появляется при наличии соответствующего электрического поля. Тогда потенциал есть

$$P(x, y) = fx + gy,$$

что с точностью до несущественной аддитивной постоянной совпадает с (26). Завершает аналогию введение идеальной твердой поверхности $y = y^0$, ограничивающей движение материальной частицы сверху; тогда (21) примет вид (25) (с учетом того, что затраты всегда считаются отрицательными).

Итак, задачи (24) и (27) полностью аналогичны и имеют совпадающие решения x^*, y^*. Заметим, что механико-экономические аналогии имеют место не только для общих формулировок проблем, но и для многих содержательных конкретных понятий (сила — предельная прибыль, реакция связей — предельные издержки и т. д.).

На плане x^*, y^* (*оптимальном плане*) максимального значения достигает не только экономический аналог потенциала — прибыль, но также еще одна величина, которой ставится в соответствие уже не механическое, а термодинамическое понятие, — энтропия. Известно, что предоставленная самой себе термодинамическая система, например, газ в изолированном сосуде (см. п. 3 § 3 гл. III), с наибольшей вероятностью переходит в состояние с наименьшей упорядоченностью параметров, характеризующих составляющие ее частицы (атомы, молекулы). В этом состоянии система находится в равновесии, ее параметры одни и те же во всех ее точках. Поэтому не существует какого-либо способа различить (упорядочить) ее части друг относительно друга: достигается наибольший (в сравнении с другими возможными состояниями) беспорядок — «хаос». Мерой этого беспорядка служит энтропия, являющаяся функцией состояния системы и принимающая максимальное значение, когда система находится в равновесии.

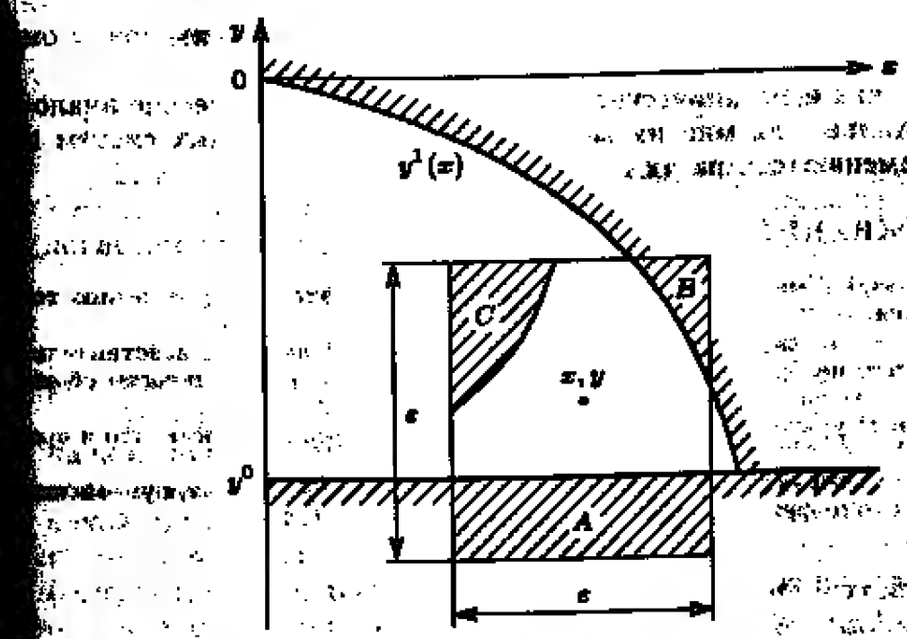


Рис. 44

Для некоторого произвольного неоптимального плана $(x, y) \neq (x^*, y^*)$ из задачи (27) рассмотрим ε-окрестность всех соседних планов $(\bar{x}, \bar{y})$ (т. е. выполнено $|\bar{x} - x| < \varepsilon$, $|\bar{y} - y| < \varepsilon$). Поскольку план (x, y) неоптимален, то в его ε-окрестности всегда существует непустое множество планов, которые: либо не удовлетворяют ресурсным ограничениям (области A и B на рис. 44); либо не имеют преимущества по прибыли в сравнении с планом (x, y) (область C на рис. 44); либо обладают обоими этими свойствами. Это означает, что в окрестно-

сти точки (x, y) можно провести частичное упорядочение планов друг относительно друга, так как известна процедура предпочтения одних планов другим. Объем множества этих «плохих» планов V зависит от x, y, ε и, очевидно, всегда меньше обема ε-окрестности:

$$V(x, y, \varepsilon) < V_\varepsilon = \varepsilon^2.$$

Рассмотрим теперь ε-окрестность оптимального плана (x^*, y^*). В отличие от случая плана (x, y) все планы из этой области менее предпочтительны по отношению к (x^*, y^*) (либо по ресурсным ограничениям, либо по величине прибыли, имеющей максимум в точке (x^*, y^*)). В этом случае объем «плохих» планов равен объему всей ε-окрестности:

$$V(x^*, y^*, \varepsilon) = V_\varepsilon.$$

Введем функцию

$$E(x, y, \varepsilon) = \frac{V(x, y, \varepsilon)}{V_\varepsilon},$$

которая служит мерой упорядоченности различных состояний данной системы. Чем «ближе» точка (x, y) к «равновесному» оптимальному состоянию (x^*, y^*), тем больше значение функции E. Очевидно, что $E < 1$ при всех $(x, y) \neq (x^*, y^*)$ и достигает своего максимального значения $E = 1$ в точке (x^*, y^*). Она и является экономическим аналогом энтропии в рассматриваемой задаче.

Так же, как и механические аналогии, термодинамические аналогии справедливы для многих весьма общих экономических систем и широко применяются при их исследовании.

УПРАЖНЕНИЯ

1. Используя замену вида $\rho(x, t) = \varphi(t) \cdot u(x, t)$, сведите (3) к уравнению теплопроводности.

2. Убедитесь в том, что квадратное уравнение (6) всегда имеет действительные корни и что необходимым и достаточным условием отрицательности обоих корней является условие $c > 0$.

3. Получите уравнения (16), (17), используя те же предположения, что и при выводе (15).

4. Дифференцируя интеграл в правой части (20) по t, x и используя уравнение (18), убедитесь, что функция $u(t, x)$ — решение уравнения (18).

§ 2. Некоторые модели финансовых и экономических процессов

Рассмотрим модели рекламной кампании, процедуры погашения взаимных долгов предприятий, простые макромодели равновесия и роста экономической системы. Обсудим роль аналогий, используемых при построении моделей, и некоторые выводы, следующие из их анализа.

1. Организация рекламной кампании. Фирма начинает рекламировать новый товар или услугу. Разумеется, что прибыль от будущих продаж должна с лихвой покрывать издержки на дорогостоящую кампанию. Ясно, что вначале расходы могут превышать

прибыль, поскольку лишь малая часть потенциальных покупателей будет информирована о новинке. Затем, при увеличении числа продаж, уже возможно рассчитывать на заметную прибыль, и, наконец, наступит момент, когда рынок насытится, и рекламировать товар далее станет бессмысленно.

Модель рекламной кампании основывается на следующих основных предположениях. Считается, что величина dN/dt — скорость изменения со временем числа потребителей, узнавших о товаре и готовых купить его (t — время, прошедшее с начала рекламной кампании, $N(t)$ — число уже информированных клиентов), — пропорциональна числу покупателей, еще не знающих о нем, т. е. величине $\alpha_1(t)\,(N_0 - N(t))$, где N_0 — общее число потенциальных платежеспособных покупателей, $\alpha_1(t) > 0$ характеризует интенсивность рекламной кампании (фактически определяемую затратами на рекламу в данный момент времени). Предполагается также, что узнавшие о товаре потребители тем или иным образом распространяют полученную информацию среди неосведомленных, выступая как бы дополнительными рекламными «агентами» фирмы. Их вклад равен величине $\alpha_2(t)\,N(t)\,(N_0 - N(t))$ и тем больше, чем больше число агентов. Величина $\alpha_2(t) > 0$ характеризует степень общения покупателей между собой (она может быть установлена, например, с помощью опросов).

В итоге получаем уравнение

$$\frac{dN}{dt} = [\alpha_1(t) + \alpha_2(t)\,N(t)]\,(N_0 - N). \qquad (1)$$

При $\alpha_1(t) \gg \alpha_2 N(t)$ из (1) получается модель типа модели Мальтуса (10) из § 1 гл. I, при противоположном неравенстве — уравнение логистической кривой (см. (12) в § 1 гл. I)

$$\frac{dN}{d\tau} = N\,(N_0 - N), \quad d\tau = \alpha_2(t)\,dt,$$

решение которого изучено в § 1 гл. I (см. также рис. 7).

Полученная аналогия вполне понятна, так как при построении данной модели и модели роста численности популяции использовалась одна и та же идея «насыщения»: скорость роста со временем какой-либо величины пропорциональна произведению текущего значения этой величины $N(t)$ на разность $N_0 - N(t)$ между ее равновесным (популяция) либо предельным (покупатели) и текущим значениями.

Аналогия между обоими процессами заканчивается, если в какой-то момент времени величина $\alpha_1 + \alpha_2 N$ становится нулевой или даже отрицательной (для этого необходимо, чтобы один или оба коэффициента $\alpha_1(t)$, $\alpha_2(t)$ стали отрицательными). Подобный негативный эффект довольно часто встречается в рекламных кампаниях различного рода и должен побудить их организаторов либо изменить характер рекламы, либо вовсе отказаться от дальнейшей пропаганды. Мероприятия по увеличению популярности товара могут, в зависимости от значений величин $\alpha_1(t)$, $\alpha_2(t)$, $N(t)$, направляться на улучшение результатов как прямой (параметр α_1), так и косвенной (параметр α_2) рекламы.

Модель (1) лишена очевидного недостатка, присущего логистическому уравнению. Действительно, оно не имеет решений, обращающихся в нуль в конечный момент времени (из соответствующей формулы для $N(t)$ в п. 5 § 1 гл. I следует, что $N(t) \to 0$ при $t \to -\infty$). Применительно к рекламе это означало бы, что часть покупателей еще до начала кампании уже знает о новом товаре. Если же рассмотреть модель (1) в окрестности точки $N(t=0) = N(0) = 0$ ($t = 0$ — момент начала кампании), считая, что $N \ll N_0$, $\alpha_2(t) N \ll \alpha_1(t)$, то уравнение (1) принимает вид

$$\frac{dN}{dt} = \alpha_1(t) N_0$$

и имеет решение

$$N(t) = N_0 \int_0^t \alpha_1(t)\, dt, \tag{2}$$

удовлетворяющее естественному начальному условию при $t = 0$.

Из (2) относительно легко вывести соотношение между рекламными издержками и прибылью в самом начале кампании. Обозначим через p величину прибыли от единичной продажи, какой бы она была без затрат на рекламу. Считаем для простоты, что каждый покупатель приобретает лишь одну единицу товара. Коэффициент $\alpha_1(t)$ по своему смыслу — число равнозначных рекламных действий в единицу времени (например, расклейка одинаковых афиш). Через s обозначим стоимость элементарного акта рекламы. Тогда суммарная прибыль есть

$$P = pN(t) = pN_0 \int_0^t \alpha_1(t)\, dt, \tag{3}$$

а произведенные затраты

$$S = s \int_0^t \alpha_1(t)\, dt.$$

Прибыль превосходит издержки при условии $pN_0 > s$, и если реклама действенна и недорога, а рынок достаточно емок, то выигрыш достигается с первых же моментов кампании (в реальности между оплатой рекламы, рекламным действием и последующей покупкой имеет место так называемый *лаг* — временная задержка, которая может быть учтена в более полных моделях). При не слишком эффективной или дорогой рекламе фирма на первых шагах несет убытки. Однако это обстоятельство, вообще говоря, не может служить основанием для прекращения рекламы. Действительно, выражение (3) и полученное с его помощью условие $pN_0 > s$ справедливы лишь при малых значениях $N(t)$, когда функции P и S растут со временем по одинаковым законам.

При увеличении $N(t)$ отброшенные в (1) члены становятся заметными, в частности усиливается действие косвенной рекламы. Поэтому функция $N(t)$ может стать более «быстрой» функцией времени, чем в формуле (3). Этот нелинейный эффект в изменении величины $N(t)$ при неизменном темпе роста издержек дает возможность скомпенсировать финансовую неудачу начальной стадии кампании.

Поясним данное утверждение в частном случае уравнения (1) с постоянными коэффициентами α_1, α_2. Заменой

$$\bar{N} = \alpha_1/\alpha_2 + N$$

оно сводится к логистическому уравнению

$$\frac{d\bar{N}}{dt} = \alpha_2 \bar{N}(\bar{N}_0 - \bar{N}), \quad \bar{N}_0 = \frac{\alpha_1}{\alpha_2} + N_0, \tag{4}$$

имеющему решение

$$\bar{N}(t) = [1 + (\bar{N}_0\alpha_2/\alpha_1 - 1)\exp(-\bar{N}_0\alpha_2 t)]^{-1}. \tag{5}$$

При этом $\bar{N}_0 = \alpha_1/\alpha_2$, так что $N(0) = 0$, и начальное условие выполняется. Из (4) видно, что производная функции $\bar{N}(t)$ и, следовательно, функции $N(t)$ может при $t > 0$ быть больше ее начального значения (при условии $\bar{N}_0 > 2\alpha_1/\alpha_2$ или $N_0 > \alpha_1/\alpha_2$). Максимум производной достигается при $\bar{N} = \bar{N}_0/2$, $N = (\alpha_1/\alpha_2 + N_0)/2$:

$$\left(\frac{d\bar{N}}{dt}\right)_m = \left(\frac{dN}{dt}\right)_m = \alpha_2 \frac{\bar{N}_0^2}{4} = \alpha_2 \frac{(\alpha_1/\alpha_2 + N_0)^2}{4}.$$

В этот период для текущей, т. е. получаемой в единицу времени прибыли имеем

$$P_m = p\frac{dN}{dt} = p\alpha_2 \frac{(\alpha_1/\alpha_2 + N_0)^2}{4}.$$

Вычитая из P_m начальную текущую прибыль $P_0 = p(dN/dt)_{t=0} = \alpha_1 N_0$ (см. (2)), получаем

$$P_m - P_0 = p\frac{(\alpha_1/\sqrt{\alpha_2} - \sqrt{\alpha_2}N_0)^2}{4},$$

т. е. разница между начальной и максимальной текущей прибылью может быть весьма значительной (см. также упр. 1). Суммарный экономический эффект от кампании (его необходимым условием является, очевидно, выполнение неравенства $P_m = p(\alpha_1/\sqrt{\alpha_2} + \sqrt{\alpha_2}N_0)^2/4 > \alpha_1 s$) определяется всем ее ходом, характеристики которого вычисляются из (4), (5) с помощью квадратуры (см. также упр. 2).

Как следует из (4), начиная с некоторого момента, продолжать рекламу становится невыгодно. Действительно, при $\bar{N}(t)$, близких к $\bar{N}_0$, уравнение (4) записывается в виде

$$\frac{d\bar{N}}{dt} = \alpha_2 \bar{N}_0(\bar{N}_0 - \bar{N}). \tag{6}$$

Его решение стремится при $t \to \infty$ к предельному значению $\bar{N}_0$ (а функция $N(t)$ — к N_0) по медленному экспоненциальному закону (упр. 2). В единицу времени появляется ничтожно малое число новых покупателей, и поступающая прибыль при любых условиях не может покрыть продолжающихся издержек.

Аналогичные характеристики вычисляются для уравнения (1) и различных его обобщений, широко используемых также для описания внедрения технологических и иных новшеств.

2. Взаимозачет долгов предприятий. Любая значительная по своим масштабам экономическая система включает в себя десятки тысяч предприятий (фирм, корпораций и т. д.), обменивающихся между собой товарами и услугами. Даже мелкая фирма, имеющая относительно небольшое число непосредственных партнеров, косвенно связана (через связи прямых и вторичных партнеров) с огромным числом предприятий, входящих в систему, и ее благополучие напрямую зависит от их состояния. Это утверждение тем более справедливо для крупных корпораций, поддерживающих партнерство с сотнями и тысячами экономических агентов.

Взаимосвязанность всех звеньев экономической системы хорошо проявляется при проведении расчетов между предприятиями за поставленную продукцию. Действительно, получив от своих клиентов выручку за предоставленный товар, предприятие тратит ее на закупку сырья и машин у других фирм, на зарплату (т. е. на покупку рабочей силы), на рекламу и на иные действия, необходимые для своего нормального функционирования. Тем самым в экономический оборот дополнительно вовлекается большое число партнеров данного предприятия. В свою очередь клиенты, получившие от предприятия товар, направляют его на дальнейшую перепродажу, используют для производства собственной продукции и т. д., также увеличивая число участвующих в экономической деятельности агентов.

Если поставки и платежи за них осуществляются своевременно (в идеале мгновенно), то с точки зрения финансового обращения экономической системе ничего не грозит. Для продолжения своей деятельности предприятиям нет необходимости ни использовать значительную часть своих финансовых ресурсов, находящихся на банковских счетах, ни, тем более, продавать основные фонды (землю, здания, оборудование, технологии). В реальности между поставкой товара и его оплатой (либо предоплатой всего товара или его части и следующей за ней поставкой) всегда существует задержка во времени. Ее минимальное значение определяется чисто техническими причинами, так как всегда требуется время на транспортировку и расфасовку товара, на проведение банковских переводов и т. д. При небольших лагах небольших партий товара (или небольших денежных сумм) предприятия на короткое время привлекают небольшую часть своих свободных финансов и затем быстро пополняют эти затраты из получаемых от своих партнеров платежей.

Однако возможны ситуации, когда по каким-либо экономическим, финансовым, внутри- или внешнеполитическим, социальным, психологическим и иным причинам время задержки платежей (поставок) становится сравнимым со временем оборота финансов, а абсолютное значение (объем) невыполненных платежей или поставок — сопоставимыми с объемом свободных оборотных средств предприятий. В этом случае возникает так называемый *кризис неплатежей*, способный привести к серьезному кризису всей экономической системы.

В самом деле, предприятие, не получившее деньги за поставленную продукцию (или оплатившее товар, но не получившее его), не может расплатиться со своими поставщиками (поскольку объем долгов предприятию сравним с величиной его свободных средств, то их использование не может принципиально улучшить ситуацию). В свою очередь поставщики не расплачиваются со своими клиентами, те — со своими и т. д. Возникают длинные цепочки неплатежей, пронизывающие всю систему. Они, очевидно, могут состоять из N звеньев, а их общее число достигать величины порядка $N!$ (N — общее число предприятий). Сумма абсолютных величин долгов по всем цепочкам может не только превышать свободные средства предприятий, но и становиться сопоставимой со стоимостью их основных фондов (речь идет именно о сумме абсолютных долгов, так как любое предприятие одновременно может быть как должником, так и кредитором своих партнеров). Система заходит в тупик — предприятия либо должны прекратить производство, либо снова занять друг у друга средства, увеличив суммарный взаимный долг.

В принципе для разрешения ситуации возможен подход, когда некоторое уполномоченное учреждение (например, главный национальный банк) выдает всем предприятиям единовременный кредит, равный сумме всех долгов. Тогда они расплачиваются между собой и затем возвращают кредит. Однако такая кредитная эмиссия может спровоцировать сильную инфляцию (производство товаров не увеличилось, а денег в обороте вдруг стало значительно больше) со всеми ее отрицательными последствиями.

В любом кризисе неплатежей определенную роль всегда играет чисто «техническая» компонента, связанная с несовершенством самой процедуры расчетов. В дальнейшем будем рассматривать кризисы, порождаемые именно этими факторами, отвлекаясь от экономических, политических и других причин их возникновения.

Сначала поясним суть проблемы на простом числовом примере для системы из трех предприятий, каждое из которых имеет свободные средства, равные условно одной финансовой единице, и основные фонды, равные 10 единицам. Пусть первое предприятие должно второму 100 единиц, второе должно третьему 100 единиц, и, наконец, третье первому также 100 единиц. Суммарный абсолютный долг предприятий равен 600 единицам и огромен в сравнении с их фондами (30 единиц), не говоря уже о свободных средствах (3 единицы). В то же время финансовое положение этой системы фактически благополучное, так как суммарный «долг» каждого предприятия в отдельности (т. е. сумма средств, которые предприятие должно другим, и другие должны ему) равен нулю. Очевидная процедура взаимозачета состоит в одномоментном аннулировании (погашении) всех долгов: объявляется, что никто никому не должен, и партнеры продолжают свою работу, буду-

чи свободными от долгового бремени. Централизованный кредит при этом, естественно, вообще не требуется.

Подобную операцию, произведенную «вручную», нельзя, конечно, реализовать для большого числа предприятий с огромным количеством финансовых обязательств. Требуются более глубокие подходы, для рассмотрения которых необходимо, прежде всего, формализовать задачу.

Итак, пусть экономическая система состоит из N предприятий, могущих иметь взаимные долги. Обозначим долги n-го предприятия m-му через x_{nm}, где $1 \leqslant n, m \leqslant N$ ($x_{nm} < 0$, если первое предприятие должно второму, и $x_{nm} > 0$ в обратном случае). Ясно, что

$$x_{nm} = -x_{mn}, \qquad x_{nn} = 0,$$

т. е. совокупность долгов описывается кососимметричной матрицей размера $N \times N$ с нулевой диагональю ($x_{nn} = 0$, поскольку предприятие самому себе должно быть не может).

Сумма всех взаимных долгов вычисляется через индивидуальные долги по простой формуле

$$X = \sum_{n=1}^{N} \sum_{m=1}^{N} |x_{nm}|. \tag{7}$$

Величина (7) служит одной из интегральных количественных характеристик финансового положения системы: если она сопоставима с суммой всех свободных средств предприятий X_0, т. е.

$$X \geqslant X_0 = \sum_{n=1}^{N} x_n, \tag{8}$$

то описываемая неравенством (8) ситуация и означает кризис неплатежей (здесь $x_n \geqslant 0$ — индивидуальные свободные средства предприятий).

Еще одна важная характеристика — баланс кредитов и долгов (сальдо) каждого предприятия

$$S_n = \sum_{m=1}^{N} x_{nm}, \tag{9}$$

причем, как очевидно из (9), возможны варианты $S_n > 0$, $S_n < 0$, $S_n = 0$. При $S_n > 0$ предприятие является в некотором смысле кредитором; предприятий-должников, т. е. тех, у кого $S_n < 0$ (при $S_n = 0$ предприятие в отношении долгов «нейтрально»). При $|S_n| < x_n$ индивидуальное финансовое состояние предприятия по существу нормальное, поскольку его реальные суммарные долги (или кредиты, «данные» им другим) меньше его свободных средств.

Аналогично, суммарное абсолютное сальдо системы

$$S = \sum_{n=1}^{N} |S_n| \tag{10}$$

служит макропоказателем ее возможного финансового «здоровья». Если $S < X_0$, то свободных средств в системе больше, чем действительных долгов, и потенциально она может успешно функционировать (подобно системе трех предприятий из приведенного выше примера).

Между величинами X и S всегда существует определенное соотношение. Для любой произвольной матрицы долгов выполняется неравенство

$$X \geqslant S, \tag{11}$$

т. е. суммарный долг никак не может быть меньше суммарного сальдо.

Задача погашения взаимных долгов состоит в том, чтобы, зная матрицу x_{nm}, найти матрицу x'_{nm} «новых» долгов, для которой выполнялось бы $X' < X$. Очевидно, что идеальным ее решением было бы $X' = S$, т. е. когда неравенство (11) становится равенством. Заметим, что тогда для благополучной по существу системы с $S \leqslant X_0$ достигалось бы соотношение $X' = S \leqslant X_0$, и после взаимозачета она могла бы нормально работать (хотя уменьшение величины X в любом случае полезно).

При построении математической модели процедуры взаимозачета долгов последовательно используется ряд действий, аналогичных проводимым при исследовании естественнонаучных объектов. Первое из них — отказ на определенном этапе от детального рассмотрения множества индивидуальных долгов и соответствующих связей между предприятиями. Переход с микроуровня на макроуровень подобен тому, как при описании большого числа частиц газа отказываются от необходимости прослеживать траекторию каждой частицы и вводят некоторые усредненные характеристики, знания которых, однако, вполне достаточно для получения подробной картины поведения объекта (см., например, вывод уравнения Больцмана в § 3 гл. III). Процедура прослеживания цепочек неплатежей, примененная

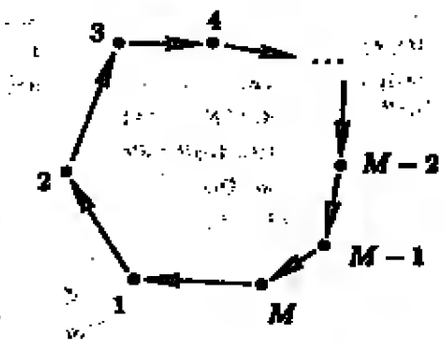


Рис. 45

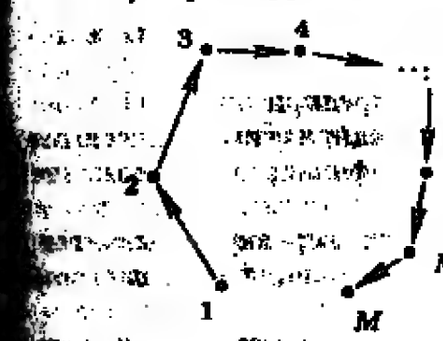


Рис. 46

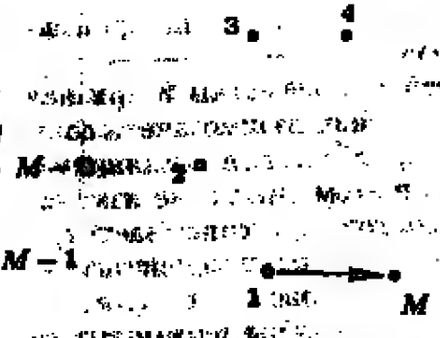


Рис. 47

выше для трех предприятий, не только трудновыполнима для N предприятий, но имеет также и принципиальный недостаток. Действительно, рассмотрим сначала цепочку, в которой каждое предприятие с

первого по M-е ($M \leqslant N$) должно другому одинаковую сумму, и такую же сумму должно M-е предприятие первому (рис. 45). Цепочка замкнута, и решение очевидно — все долги в цепочке погашаются. Пусть теперь M-е предприятие не должно первому (рис. 46). Тогда цепочка разомкнута, и этот метод неприменим. В то же время простое решение заключается в том, что долги предприятий со второго по $(M-1)$-е аннулируются, а долг первого переадресовывается M-му (рис. 47). Экономический смысл переадресации соответствует вексельному обращению, когда долговое обязательство меняет своих хозяев, и в результате у должника (первое предприятие) появляется новый кредитор (M-е предприятие).

В отличие от ситуации с долгами в цепочках полная система долгов по всем цепочкам замкнута, так как рассматриваются взаимные долги. В самом деле, из свойства $x_{nm} = -x_{mn}$ следует, что

$$\sum_{n=1}^{N}\sum_{m=1}^{N} x_{nm} = 0$$

для любой совокупности неплатежей. Учитывая, что $S_n = \sum_{m=1}^{N} x_{nm}$, из последнего равенства получаем

$$\sum_{n=1}^{N} S_n = 0, \tag{12}$$

или

$$\sum_{S_n>0} S_n = -\sum_{S_n<0} S_n = \frac{S}{2}, \tag{13}$$

т. е. сумма положительных сальдо предприятий равна по абсолютной величине сумме отрицательных сальдо. Рассматриваемая на макроуровне система взаимных долгов обладает свойством «симметричной консервативности» (13), а «закон сохранения» (12) — аналог обычных законов сохранения (массы, энергии и т. д.) применительно к изучаемой ситуации.

Равенство (13) проясняет построение математической модели идеального взаимозачета, который производится при следующих естественных условиях:

1) все долги x_{nm} известны и признаются предприятиями;

2) при проведении взаимозачета сальдо предприятий S_n остаются неизменными: $S'_n = S_n$, т. е. индивидуальное финансовое положение каждого из них в этом смысле не изменяется;

3) часть долгов x_{nm} списывается, а часть переадресовывается, т. е. у предприятий могут появиться новые должники и кредиторы и исчезнуть часть старых.

Суть макропроцедуры взаимозачета состоит в том, что вместо величин x_{nm} рассматриваются величины S_n. Предприятия с $S_n < 0$ объявляются должниками (в размере своих сальдо), а предприятия с $S_n > 0$ — кредиторами (в тех же размерах). Затем долги предприятий с $S_n < 0$ каким-то образом распределяются между кредиторами,

т. е. находится новая система долгов x'_{nm}. При этом выполнены закон сохранения (12), условие 2) и достигается равенство $X' = S$; поэтому решение задачи является оптимальным.

Таких оптимальных решений может быть, вообще говоря, много, так как распределять долги между кредиторами можно разными способами. Приведем два наиболее простых и наглядных. Первое из них дается несложной формулой, по которой новые долги вычисляются через старые:

$$x'_{nm} = \frac{S_n\,|S_m| - S_m\,|S_n|}{S}. \tag{14}$$

Согласно алгоритму (14) долг любого предприятия (равный S_n, если $S_n < 0$) расписывается по предприятиям-кредиторам в долях, пропорциональных величинам их сальдо (равным S_m, если $S_m > 0$). Предприятиям с бóльшим положительным сальдо причитается от каждого из должников бóльшая часть его долгов, причем в сумме они дают величину S_m (упр. 5). Для предприятий с нулевым сальдо взаимозачет сводится к погашению всех их долгов и всех долгов им.

Заметим, что в решении (14) для новых долгов имеем $x'_{nm} = 0$ при $S_n < 0$, $S_m < 0$ либо при $S_n > 0$, $S_m > 0$ (после взаимозачета должники не должны должникам, а кредиторы — кредиторам). Это означает, что число получившихся финансовых связей между предприятиями значительно меньше максимально возможного, когда каждое предприятие является должником или кредитором любого другого, и матрица долгов не имеет нулевых элементов (кроме, разумеется, диагональных).

Количество связей может быть значительно уменьшено, если провести предварительное упорядочивание предприятий по абсолютным значениям их сальдо и установить непосредственные связи между должниками и кредиторами одного масштаба (крупных с крупными, мелких с мелкими и т. д.). Эта процедура допускает простую геометрическую интепретацию. На рис. 48 на верхней прямой линии описано распределение сальдо кредиторов (в убывающем порядке). Длина отрезков этой прямой равна величине сальдо каждого предприятия $S_p > 0$, $1 < p < N$, а ее общая длина, очевидно, равна $S/2$. На нижней прямой описано распределение сальдо должников $S_q < 0$, $1 < q < N$, $p + q \leqslant N$ (сальдо взяты с обратным знаком) также в убывающем порядке. Ее длина согласно (13) также равна $S/2$. Штриховые линии, проведенные из узлов нижней прямой, делят «прямую кредиторов» на q отрезков, равных величине долга каждого предприятия. Этот долг либо достается одному кредитору, либо делится между несколькими в соответствии с расположением узлов верхней прямой относительного данного отрезка.

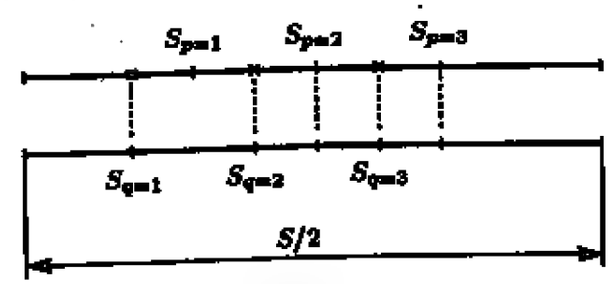


Рис. 48

Описанный алгоритм оптимален по критерию $X' = S$ и представляется наилучшим по числу связей, остающихся после взаимозачета.

Таблица 2

Пример взаимозачета в системе с $N = 10$ и начальной матрицей долгов с 90 ненулевыми недиагональными элементами

	1	2	3	4	5	6	7	8	9
Начальная матрица ($X = 3729$)									
2	−25								
3	−1	−20							
4	4	25	−2						
5	25	−450	25	30					
6	−15	150	−30	20	−926				
7	3	−40	3	3	5	25			
8	1	−22	−2	−2	4	−15	5		
9	10	322	−15	−25	496	−800	−10	20	
10	1	−25	−2	1	−20	15	−1	−3	30
Конечная матрица ($X' = S = 62$)									
2	2								
3	0	0							
4	0	0	0						
5	0	0	0	0					
6	0	0	0	0	−28				
7	1	0	0	0	0	0			
8	0	−7	0	0	0	0	0		
9	0	−18	0	0	−2	0	0	0	
10	0	0	0	0	0	4	0	0	0

Пример подобного взаимозачета в системе с $N = 10$ и начальной матрицей долгов с 90 ненулевыми недиагональными элементами приведен в табл. 2. Конечная матрица содержит лишь 14 ненулевых элементов. В специальных случаях у одного должника остается один кредитор, и наоборот (упр. 6).

Отметим, что эти и другие процедуры взаимозачета имеют смысл лишь при выполнении условий 1)–3), т. е. при определенном соглашении между предприятиями. Причины, не позволяющие придерживаться данного соглашения, могут быть весьма разнообразными — от нежелания платить долги потому, что это выгодно должнику, до последствий санкций международных или иных организаций, когда финансовые средства предприятий замораживаются. Эти обстоятельства и определяют рамки применимости модели взаимозачета, при построении которой существенно использовались аналогии с моделями некоторых естественнонаучных объектов.

3. Макромодель равновесия рыночной экономики. Любой участник рыночного экономического процесса действует в соответствии со своими индивидуальными интересами (извлечение прибыли, улучшение условий труда, минимизация риска, экономия ресурсов и т. д.). Простейший вариант такой системы — экономика с совершенной конкуренцией, когда каждый субъект экономически ничтожно мал и не оказывает непосредственного влияния на уровень производства, цены,

зарплату и другие макропоказатели. В то же время разобщенные действия экономических агентов могут складываться через существующие в системе отношения купли-продажи в совокупную согласованную картину действий работодателей и наемных рабочих, финансистов и вкладчиков и т. д.

Если в результате такого коллективного взаимодействия общее производство товаров и услуг в системе согласовано с общим спросом на них, то такое состояние экономики называется *равновесным*, а устанавливающиеся при этом цены — *равновесными рыночными ценами*. Баланс между спросом и предложением имеет место, как подразумевается, не при произвольных, а именно при этих рыночных ценах, что означает, в частности, платежеспособность спроса.

Одна из важных задач экономической науки — определение условий равновесия экономики, и в том числе равновесных рыночных цен. Наиболее простые математические модели экономического равновесия строятся при следующих предположениях:

1) совершенная рыночная конкуренция, означающая отсутствие как крупных производственных корпораций (и, тем более, монополий), так и объединений работников, могущих диктовать свои условия для всей системы;

2) неизменность производственных возможностей системы: оборудование, производственные помещения, технологии не изменяются со временем;

3) неизменные во времени экономические интересы партнеров: предприниматели не пытаются увеличить свою прибыль, рабочие — зарплату, инвесторов устраивают проценты, получаемые по ценным бумагам, и т. д.

Отвечающие этим предположениям модели описывают весьма частный случай «застывшей» во времени идеальной рыночной экономики. Однако они дают ответ на вопрос о возможности существования экономического равновесия, формирующегося из рыночного «хаоса», и, кроме того, связывают между собой основные макропоказатели экономической системы.

Одна из таких макромоделей — *модель Кейнса* — рассматривает в качестве агентов нанимателей и нанимаемых, потребителей и сберегателей, производителей и инвесторов, действующих на рынках рабочей силы, продуктов и денег, т. е. распределяющих и обменивающих эти товары (труд, продукты, деньги) между собой.

Первый макропоказатель системы — *национальный доход* Y, являющийся единственным (для простоты) продуктом, производимым ею в единицу времени. Этот продукт вырабатывается производственным сектором экономики, а его величина дается функцией F, зависящей от количества и качества ресурсов, состава основных фондов и *числа занятых работников* R (второй макропоказатель). В соответствии с предположением 2) в состоянии равновесия производственная функция R, а с нею и продукт Y определяется лишь занятостью, т. е.

$$Y = F(R). \tag{15}$$

Относительно $F(R)$ обычно считается, что $F(0) = 0$, $F'(R) > 0$, $R > 0$ и $F''(R) < 0$ при $R > 0$ (рис. 49). Функция $F(R)$ обладает свойством «насыщения»: с ростом R выпуск растет все медленнее. Такой подход вполне оправдан, поскольку при излишне большом числе занятых на производстве для них попросту не найдется соответствующего фронта работ. Один или несколько старателей, обнаруживших золотую жилу, быстро и без помех достигнут своей максимальной производительности; при большом количестве рабочих они начинают мешать друг другу, и их индивидуальная производительность уменьшается; наконец, при очень большом числе старателей добыча золота вообще перестает расти, так как вновь прибывающие не могут подступиться к месту разработки.

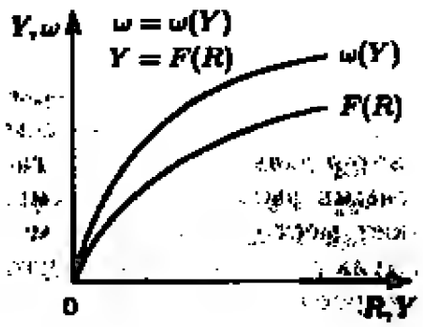


Рис. 49

Соотношение (15) дает связь между рынками труда (R) и продукта (Y). Дополнительное соотношение определяется с помощью одного из основных постулатов классической политэкономии:

4) *заработная плата* s работника равна стоимости продукта, которая была бы потеряна при уменьшении занятости на одну единицу (заработная плата равна предельному продукту труда).

Заметим, что в постулате 4) не учитываются (считаются малыми) другие издержки, которые отпали бы в результате сокращения одного рабочего места (затраты на ресурсы, оборудование и т. д.). Таким образом, из постулата получаем

$$\Delta Y^{(1)} \cdot p = s,$$

где $\Delta Y^{(1)}$ — количество продукта, потерянное при уменьшении занятости на единицу, p — *цена продукта* (так что слева в данном равенстве записана величина потерянной стоимости). Если занятость изменилась на величину ΔR, то из последнего равенства, очевидно, имеем

$$\Delta Y \cdot p = s \cdot \Delta R,$$

где $\Delta Y = \Delta Y^{(1)} \Delta R$ — стоимость, потерянная или полученная при изменении числа работников на ΔR. Считая ΔR и ΔY малыми в сравнении с R и Y, перепишем последнее равенство в дифференциальной форме:

$$\frac{dY}{dR} = \frac{s}{p},$$

или, принимая во внимание (15),

$$F'(R) = \frac{s}{p}. \qquad (16)$$

Поскольку $F(R)$ задана (а с нею известна функция $F'(R)$), то при известных макропоказателях s и p из (16) можно найти уровень занятости R, а из (15) и величину продукта Y. Напомним: этот уровень отвечает числу работников, согласных трудиться за данную зарплату при данных ценах и других характеристиках системы, а не вообще возможному числу наемных рабочих. Предполагается, что для обеспечения равновесного уровня занятости всегда найдется достаточное количество желающих работать на существующих условиях, т. е.:

5) предложение труда не сдерживает производства, число занятых определяется спросом на труд со стороны предпринимателей.

Два уравнения (15), (16) содержат четыре величины. Относительно одной из них предполагается, что:

6) заработная плата s в модели считается заданной.

Она определяется в результате компромисса между работодателями и нанимаемыми (реальная же зарплата зависит также и от уровня цен).

Очевидно, для построения замкнутой модели необходимо дальнейшее рассмотрение рынка продукта и рынка финансов. Произведенный продукт частично тратится на потребление, а частично сберегается:

$$Y = S + \omega,$$

где ω — *потребляемая часть* (в экономику не возвращается), а S — сберегаемая часть, возвращающаяся в экономическую систему (или фондообразующий продукт).

Соотношение между величинами S и ω определяется из следующих соображений. Относительно величины ω считается, что:

7) потребляемая часть выпуска зависит от величины самого выпуска, т. е. $\omega = \omega(Y)$.

При этом функция $\omega(Y)$ обладает свойством «насыщения» так же, как и функция $F(R)$: чем больше выпуск, тем меньшая доля дополнительного выпуска ΔY тратится на потребление (рис. 49) и тем большая доля сберегается. Величина $d\omega/dY = c(Y)$ называется *склонностью к потреблению* и лежит в пределах $0 < c < 1$, иначе при малых выпусках потреблялось бы больше продукта, чем производилось бы (величина $d = 1 - c$ — склонность к накоплению).

Фондообразующий продукт

$$S = Y - \omega(Y) \qquad (17)$$

вкладывается инвесторами в экономику с целью получить в будущем с этих инвестиций доход. В модели считается, что инвестиции эквивалентны отложенному (отнесенному на будущее) потреблению и поэтому определяются еще одним финансовым макропоказателем системы — *нормой банковского процента* r. Действительно, сделав инвестиции в размере A и получив через год доход $D = Ar$, инвестор ничего не теряет (в данном примере и не выигрывает) по сравнению с вложением этих средств в банк под процент r. В обоих случаях сегодняшнее потребление откладывается ради возможности большего потребления в следующем году. Спрос на инвестиции задается функцией $A(r)$ такой, что $A'(r) < 0$ при $0 < r < r_1$ и $A(r) = 0$ при $r \geqslant r_1$: при большой норме процента инвестиции отсутствуют (рис. 50).

В условиях равновесия предложение фондообразующего продукта $S(Y)$ сбалансировано со спросом на инвестиции $A(r)$:

$$S(Y) = A(r),$$

или, учитывая (17),

$$Y - \omega(Y) = A(r). \qquad (18)$$

Для окончательного замыкания модели рассматривается рынок финансов. Деньги нужны экономическим агентам для покупки фондообразующего продукта, для потребления, а также как одно из средств накопления. Считается, что деньги выпускает государство, и их количество (*предложение*) Z является заданным управляющим параметром системы. Относительно спроса на деньги делается следующее предположение:

8) спрос на деньги представляет собой сумму операционного и спекулятивного спроса.

Операционный спрос определяется количеством денег, которое надо иметь на руках, чтобы производить покупки товара Y (как фондообразующего, так и идущего на потребление). Если цена продукта равна p, а *время обращения* равно τ, то операционный спрос равен, очевидно, величине τpY.

Спекулятивный спрос связан с величиной нормы процента r. Если норма процента высока, то большую часть денег их владельцы предпочитают хранить в банке, рассчитывая на хороший доход и жертвуя более высокой степенью ликвидности банкнот (способностью обмениваться на продукты) в сравнении с банковскими обязательствами. При низкой процентной ставке спекулятивный спрос увеличивается: владельцы желают иметь на руках все больше банкнот, аккумулируя в них свои накопления. Поэтому спекулятивный спрос задается функцией $I(r)$ (рис. 50) такой, что $I'(r) < 0$ при $r > r_2$ и $I(r)$ резко возрастает при $r \to r_2$ ($\lim I(r) = \infty$, $r \to r_2$; владельцы денег не приобретают обязательств банка). Естественно считать $r_2 < r_1$, так как в противном случае либо инвестиции равны нулю, и говорить об экономическом равновесии не приходится, либо функция $I(r)$ не определена, и рассмотрение не имеет смысла.

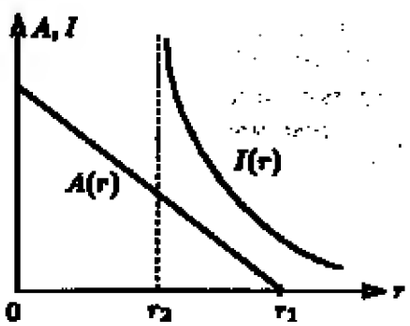


Рис. 50

Так как финансовый рынок находится в равновесии, то баланс («закон сохранения») денег в системе дается уравнением

$$Z = \tau pY + I(r). \qquad (19)$$

Сводя воедино уравнения (15), (16), (18), (19), приходим к *математической модели рыночного равновесия*, полученной в предположениях 1)–8):

$$\begin{aligned} &Y = F(R),\\ &F'(R) = s/p,\\ &Y - \omega(Y) = A(r),\\ &Z = \tau pY + I(r). \end{aligned} \qquad (20)$$

В модели (20) задаются параметры системы s (ставка заработной платы), Z (предложение денег) и технический параметр τ. Функции F, F', ω, A, I — известные функции своих аргументов с описанными выше свойствами. По этим входным данным из модели определяются четыре неизвестных величины: Y (выпуск продукта), R (занятость), p (цена продукта) и r (норма прибыли).

Исключая из (20) величины p, r, Y, уравнения (20) легко свести к одному уравнению относительно R

$$-\frac{\tau s F(R)}{F'(R)} + Z = I\{A^{-1}[F(R) - \omega(F(R))]\}, \qquad (21)$$

где A^{-1} — функция, обратная к функции A. Найдя из (21) значение R, из (20) нетрудно определить все остальные искомые величины.

Докажем с помощью нестрогих, но зато простых построений существование решения (21), основываясь на анализе графиков функций, входящих в его левую и правую части.

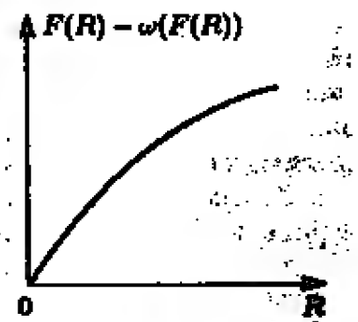


Рис. 51

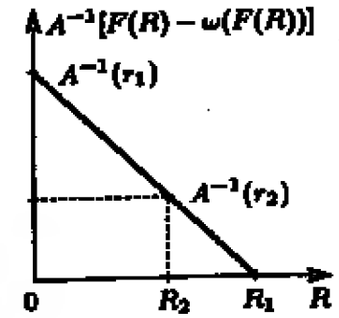


Рис. 52

Функция $F(R) - \omega(F(R))$ — монотонно растущая функция R, равная нулю при $R = 0$ (рис. 51). Ее монотонность следует из условия $d\omega(F(R))/d(F(R)) = c < 1$, а рост по мере увеличения R — из условия $dF(R)/dR > 0$. Данная функция является аргументом для монотонной функции A^{-1}, и из свойств функции A (рис. 50) легко установить качественный вид зависимости A^{-1} от R (рис. 52), причем $A^{-1} \equiv 0$ при $R > R_1$ (R_1 — некоторое значение величины R, $0 < R_1 < \infty$). В свою

очередь A^{-1} служит аргументом монотонной функции I, свойства которой таковы (рис. 50), что как функция R она имеет вид, изображенный на рис. 53 (для значений $R > R_2$ функция I не определена).

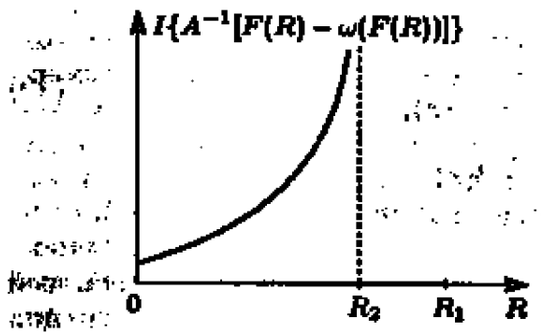


Рис. 53

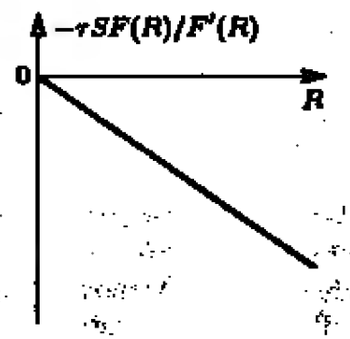


Рис. 54

Рассмотрим теперь левую часть уравнения (21). Функция $-\tau sF(R)/F'(R)$ равна нулю при $R = 0$ (считается, что $F'(0) \neq 0$); (см. рис. 49). Ее первая производная по R, как следует из свойств функций $F'(R) > 0$, $F''(R) < 0$, отрицательна, т. е. она монотонно убывает (рис. 54).

Совмещая графики левой (кривая *2*) и правой (кривая *1*) частей уравнения (21) на одном рис. 55, убеждаемся в том, что при достаточно большом значении управляющего параметра Z кривые пересекаются в некоторой точке R_0, $0 < R_0 < \infty$. Точка пересечения единственна в силу монотонности графиков. Следовательно, модель (20) действительно имеет единственное решение, описывающее равновесное состояние экономики.

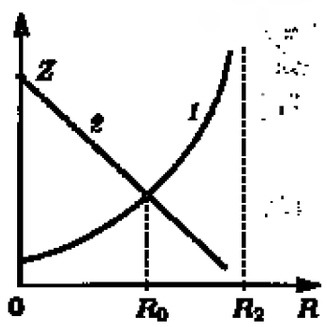


Рис. 55

Однако значение модели этим не ограничивается. Она может быть использована для сравнительного анализа различных, но близких состояний равновесия (не отвечая, естественно, на вопрос, каким образом система приходит в равновесное состояние или выходит из него). Допустим, что значения равновесных параметров s_0 и Z_0 изменились на малые величины δs_0 и δZ_0 при переходе из одного равновесного состояния в другое (параметр τ считаем неизменным). Тогда изменятся и все остальные характеристики системы. Их можно найти из (20), имея в виду, что оба сравниваемых состояния равновесны. Например, из второго уравнения (20), используя разложение в ряд Тейлора, получаем

$$\frac{s_0}{p_0}\frac{\delta p}{p} = \frac{\delta s}{p_0} - F''(R_0)\,\delta R.$$

Проведя аналогичную процедуру с остальными уравнениями системы (20) и сводя результаты воедино, получаем

$$\frac{\delta p}{p_0} = a_1(\delta A + \delta\omega) + a_2(\delta Z - \delta I) + a_3\delta s - a_4\delta Y. \qquad (22)$$

В выражении (22) присутствуют все характеристики изучаемой системы (коэффициенты $a_i < 0$, $i = 1, \ldots, 4$, определяются равновесными значениями величин s_0, p_0, Y_0, r_0, функций R, ω, A, I и их производных). Поэтому с его помощью можно анализировать весь комплекс изменений, происходящих при переходе от одного равновесного состояния к другому (так называемый *системный подход*). Пусть, например, при неизменном числе занятых (т. е. $\delta R = 0$, $\delta Y = 0$), неизменных зарплате ($\delta s = 0$) и уровне потребления ($\delta\omega = 0$) требуется понизить цену ($\delta p < 0$), т. е. увеличить реальную заработную плату работающих. Тогда необходимо стремиться уменьшить инвестиции ($\delta A < 0$), снизить общий объем денег ($\delta Z < 0$) и увеличить спекулятивный спрос на них ($\delta I > 0$). Заметим, что требования, вытекающие из анализа соотношения (22), могут быть, вообще говоря, противоречивыми.

Разумеется, эта и другие системы мероприятий, могущие следовать из построенной модели, не реализуются автоматически путем соответствующей вариации параметра Z или s (или обоих параметров). Модель (20) лишь указывает на необходимые изменения в поведении экономических агентов. Вопрос о том, как действительно обеспечить эти изменения, убедив участников рыночного процесса принять их, выходит за рамки рассматриваемой модели. Его решение связано с изучением еще более трудноформализуемых объектов. При исследовании подобных объектов широко применяются подходы, почерпнутые из естественнонаучной сферы, такие, как идея насыщения, переходы с микро- на макроуровень, использование «законов сохранения», представления о стационарности и равновесности и др.

4. Макромодель экономического роста. В растущей экономике число работающих $R(t)$ не постоянно, а увеличивается с течением времени. В простейшей модели считается, что темп прироста занятых работников пропорционален числу уже работающих:

$$\frac{dR}{dt} = \alpha R(t).$$

Поэтому $R(t) = R_0 e^{\alpha t}$ — заранее известная функция времени (величина α задается, $R_0 = R(0)$ — число работающих в начальный момент $t = 0$). Работники производят национальный доход $Y(t)$, который частично идет на потребление и частично на накопление:

$$Y(t) = \omega + A. \qquad (23)$$

Накопленная часть продукта A возвращается в экономику с тем, чтобы скомпенсировать выбывающие из строя производственные мощности, а также для создания новых мощностей. Под *мощностью* $M(t)$ подразумевается максимально возможный выпуск продукта экономикой. Реальный выпуск продукта зависит, естественно, от числа работающих и задается производственной функцией вида

$$Y(t) = M(t) \cdot f(x(t)). \qquad (24)$$

В формуле (24) величина $x(t) = R(t)/M(t)$ по своему смыслу — количество работающих на единице мощности. Относительно функции $f(x)$ делаются следующие предположения: $f(0) = 0$, $f' > 0$ (выпуск растет с увеличением числа занятых) и $f'' < 0$ (насыщение). Функ-

ция $f(x)$ определена для значений x на отрезке $0 \leqslant x \leqslant x_M$, где $x_M = R_M/M$, а $R_M(t)$ — число рабочих мест в хозяйстве при мощности $M(t)$. Если все места заполнены, то выпуск $Y(t)$ по определению равен $M(t)$, т. е. для $f(x)$ должно выполняться условие $f(x_M) = 1$.

Одна из главных задач *теории экономического роста* — нахождение оптимальных в некотором смысле способов разделения произведенного продукта на потребляемую и накапливаемую части. Критерием оптимальности можно выбрать, например, душевое потребление (количество продукта, потребляемого одним работающим), т. е. величину $c(t) = \omega(t)/R(t)$.

Сбереженный в единицу времени продукт $A(t)$ расходуется на создание *новой мощности*:

$$A(t) = a\,I(t),$$

где $a > 0$ — считающееся заданным и постоянным количество фондообразующего продукта, необходимое для создания единицы новой мощности, $I(t)$ — число единиц новой мощности.

Темп выбытия существующей мощности предполагается пропорциональным величине самой мощности, т. е. величине $\beta M(t)$, коэффициент выбытия $\beta > 0$ задается постоянным.

В итоге для изменения функции $M(t)$ получаем балансное соотношение

$$\frac{dM}{dt} = I(t) - \beta M(t). \tag{25}$$

Уравнения (23)–(25) содержат четыре неизвестных величины — $Y(t)$, $\omega(t)$, $M(t)$, $I(t)$. Для замыкания модели предположим, что скорость введения новой мощности пропорциональна величине уже существующей мощности: $I(t) = \gamma M(t)$, где $\gamma > 0$ (величина, обратная характерному времени наращивания мощности) считается заданной и постоянной (естественно, $\gamma > \beta$). Тогда решение уравнения (25) легко находится:

$$M(t) = M_0\, e^{(\gamma-\beta)t}, \tag{26}$$

а вместе с ним определяются и все остальные неизвестные величины.

Проанализируем простой, но показательный случай экономического роста, когда мощность увеличивается со временем в том же темпе, что и число работающих. Для этого, очевидно, необходимо, чтобы выполнялось равенство

$$\gamma - \beta = \alpha. \tag{27}$$

Оно означает также, что с тем же темпом растут функции $Y(t)$ (поскольку $f(x(t)) = f(x = R_0/M_0) = \mathrm{const}$) и $\omega(t)$, $I(t)$.

Найдем число работающих и соотношение между потреблением и накоплением, при которых душевое потребление работников максимально. По определению

$$c(t) = \frac{\omega(t)}{R(t)} = \frac{Y(t) - A(t)}{R(t)}.$$

Учитывая, что $Y(t) = M(t)\,f(x)$, $A(t) = a\gamma M(t)$, и принимая во вни-

мание (26), (27), получаем

$$c(t) = c = \frac{f(x) - a(\alpha+\beta)}{x}, \tag{28}$$

т. е. душевое потребление со временем не изменяется. Его максимум, как очевидно из (28), достигается при условии

$$\frac{dc}{dx} = \frac{d}{dx}\left[\frac{f(x) - a(\alpha+\beta)}{x}\right] = 0,$$

которое дает уравнение для искомой величины x_m

$$x_m f'(x_m) - f(x_m) + a(\alpha+\beta) = 0. \tag{29}$$

Это уравнение всегда имеет единственное решение $0 < x_m \leqslant x_M$ (упр. 9). Заметим, что помимо всех сделанных предположений, для реализации рассматриваемого режима экономического роста необходимо согласование численности работающих R_0 с мощностью M_0 в начальный момент времени так, чтобы $R_0/M_0 = x_m$.

Норма накопления, обеспечивающая максимальное значение c_m,

$$n_m = \frac{A_m}{Y_m},$$

находится из равенств $Y_m = M_m f(x_m)$, $A_m = a\gamma M_m$ и из (27), (29):

$$n_m = 1 - x_m \frac{f'(x_m)}{f(x_m)}, \tag{30}$$

и называется *нормой золотого правила роста Солоу*.

Если условие (27) не выполнено, то режимы роста экономики становятся более сложными, и оптимизация их характеристик проводится на всем рассматриваемом временном интервале. Напомним, что построенная модель и подобные ей модели не учитывают изменений в производственных отношениях (они считаются постоянными) и оперируют в основном технологическими связями, давая, в частности, верхние технологические ограничения на темп экономического роста (упр. 11). При их получении также широко используются аналогии с естественнонаучными объектами.

УПРАЖНЕНИЯ

1. Пусть параметры α_1, N_0, p, s фиксированы. Найдите значение параметра α_2, при котором функция $P_m(\alpha_2)$ достигает минимального значения, и убедитесь, что в этом случае текущая прибыль $P(t)$ максимальна в момент $t = 0$, т. е. на старте рекламной кампании.

2. Неравенство $P_m > \alpha_1 s$ (зависящее от $\alpha_1 > 0$ нелинейным образом) является необходимым условием прибыльности рекламы. Считая параметры α_2, N_0, p, s фиксированными, найдите области значений параметра α_1, в которых при росте α_1 данное неравенство усиливается (ослабляется).

3. Считая, что эффект насыщения от рекламы происходит при $N(t) \approx N_0$, найдите из уравнения (6) момент времени, когда продолжение кампании станет заведомо убыточным.

4. Убедитесь в справедливости неравенств (11).

5. Проверьте, используя формулы (7), (9), (10), с помощью равенства (13), что для решения (14) выполнены свойство $x'_{nm} = -x'_{mn}$, условие 2) и критерий оптимальности $X' = S$.

6. Покажите, что в соответствии с описанной процедурой начальная матрица долгов вида

$$x_{nm} = \begin{vmatrix} 0 & +1 & \dots & +1 & \dots & +1 & +1 \\ -1 & 0 & \dots & +1 & \dots & +1 & +1 \\ \dots & & \ddots & \dots & \iddots & \dots & \\ -1 & -1 & \dots & 0 & \dots & +1 & +1 \\ \dots & & \iddots & \dots & \ddots & \dots & \\ -1 & -1 & \dots & -1 & \dots & 0 & +1 \\ -1 & -1 & \dots & -1 & \dots & -1 & 0 \end{vmatrix}$$

где N четное, $X = N(N-1)$, принимает вид

$$x'_{nm} = \begin{vmatrix} 0 & 0 & \dots & 0 & 0 & \dots & 0 & N-1 \\ 0 & 0 & \dots & 0 & 0 & \dots & N-3 & 0 \\ \dots & & \ddots & \dots & \iddots & & \dots & \\ 0 & 0 & \dots & 0 & +1 & \dots & 0 & 0 \\ 0 & 0 & \dots & -1 & 0 & \dots & 0 & 0 \\ \dots & & \iddots & \dots & \ddots & & \dots & \\ 0 & -(N-3) & \dots & 0 & 0 & \dots & 0 & 0 \\ -(N-1) & 0 & \dots & 0 & 0 & \dots & 0 & 0 \end{vmatrix}$$

причем $X' = N^2/2 = S < X$.

7. Выражение $\Delta Y = \Delta A/(1-c)$, где ΔA — изменение в уровне инвестиций, а ΔY — соответствующий прирост (убыль) выпуска продукта, называется *соотношением мультипликатора Кейнса*. Получите его из модели (20).

8. Получите выражения для коэффициентов a_i, $i = 1, \dots, 4$, в (22) и убедитесь в том, что все они положительны.

9. Докажите, используя свойства функции $f(x)$, существование единственного решения уравнения (29).

10. Используя рассуждения, применявшиеся при выводе (30), покажите, что для величины максимального душевого потребления c_m справедливо равенство $c_m = f'(x_m)$ (величина $f'(x_m)$ называется *предельной производительностью труда* — это прирост выпуска продукта при увеличении занятости на единицу).

11. Покажите, что если весь произведенный продукт идет на накопление и мощности загружены полностью, то экономический рост реализуется с темпом $Y(t) = Y_0 t e^{\gamma_m t}$, $\gamma_m = 1/a - \beta$.

§ 3. Некоторые модели соперничества

Построим модели различных видов соперничества — двухвидовой борьбы в популяциях, гонки вооружений, боевых действий. Покажем общность методологических подходов, применяемых при получении и анализе этих моделей.

1. Взаимоотношения в системе «хищник—жертва». Строго говоря, эти отношения (как и отношения в аналогичной системе «паразит—хозяин») не могут называться соперничеством. «Соперничество» жертвы с хищником выражается в изменении численности жертвы, которая в свою очередь сказывается на численности хищника. Действительно, никакой организм (тем более популяция) не живет изолированно, а взаимодействует со своим окружением. Широко распространенный вид взаимодействия — использование одними живыми организмами (животными, птицами, рыбами, насекомыми) других организмов в качестве пищи.

Математическая модель наиболее простой, т. е. двухвидовой системы «хищник—жертва» основывается на следующих предположениях:

1) численности популяций жертв N и хищников M зависят только от времени (точечная модель, не учитывающая пространственное распределение популяции на занимаемой территории; ср. с моделью сообщества амеб в § 1);

2) в отсутствие взаимодействия численность видов изменяется по модели Мальтуса из п. 3 § 1 гл. I; при этом число жертв увеличивается, а число хищников падает, так как им в этом случае нечем питаться:

$$\frac{dN}{dt} = \alpha N, \quad \frac{dM}{dt} = -\beta M, \quad \alpha > 0, \quad \beta > 0;$$

3) естественная смертность жертвы и естественная рождаемость хищника считаются несущественными;

4) эффект насыщения численности обеих популяций не учитывается;

5) скорость роста численности жертвы уменьшается пропорционально численности хищников, т. е. величине cM, $c > 0$, а темп роста хищников увеличивается пропорционально численности жертвы, т. е. величине dN, $d > 0$.

Объединяя предположения 1)–5), приходим к системе *уравнений Лотки—Вольтерра*

$$\begin{aligned} \frac{dN}{dt} &= (\alpha - cM)\,N, \\ \frac{dM}{dt} &= (-\beta + dN)\,M, \end{aligned} \tag{1}$$

из которой по начальным численностям $N(0) = N(t=0)$, $M(0) = M(t=0)$ определяется численность популяции в любой момент $t > 0$.

Нелинейную систему (1) удобно исследовать в плоскости переменных N, M, для чего первое уравнение поделим на второе:

$$\frac{dN}{dM} = \frac{(\alpha - cM)\,N}{(-\beta + dN)\,M}. \tag{2}$$

Уравнения (1), (2) имеют положение равновесия (или стационарное, не зависящее от времени решение)

$$M_0 = \frac{\alpha}{c}, \quad N_0 = \frac{\beta}{d}. \tag{3}$$

Зададимся вопросом об устойчивости положения равновесия (3). Под этим подразумевается следующее. Если начальные численности в точности равны величинам (3), то как с течением времени они изменяются? Если по каким-то причинам численности ненамного отклоняются от величин M_0, N_0, то вернется ли система в положение равновесия? Наконец, если начальные значения $N(0)$, $M(0)$ заметно отличаются от равновесных, то каким образом они меняются со временем относительно величин N_0, M_0?

Чтобы понять временную динамику функций $N(t)$, $M(t)$, преобразуем уравнение (2) к виду

$$dN\,(-\beta + dN)\,M = dM\,(\alpha - cM)\,N,$$

поделим обе части получившегося равенства на величину NM и перенесем все члены в левую часть:

$$\beta\frac{dN}{N} - d\,dN + \alpha\frac{dM}{M} - c\,dM = 0. \tag{4}$$

Уравнение (4) нетрудно проинтегрировать и получить соотношение

$$\beta\ln N - dN + \alpha\ln M - cM = \text{const},$$

где константа в правой части определяется по начальным значениям N_0, M_0. Другими словами, уравнение (2), или, что то же самое, система (1) имеет интеграл вида

$$\ln N^{\beta} + \ln e^{-dN} + \ln M^{\alpha} + \ln e^{-cN} = C.$$

Потенцируя последнее выражение, получаем интеграл в виде

$$N^{\beta}e^{-dN} = C_1 M^{-\alpha}e^{cM}, \qquad C_1 > 0. \tag{5}$$

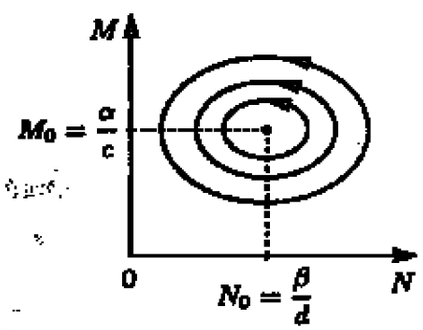


Рис. 56

Существование интеграла (5) дает возможность ответить на поставленные вопросы (на рис. 56 изображены фазовые траектории системы (1); стрелками показано направление движения по траекториям с течением времени):

а) если $N(0) = N_0$, $M(0) = M_0$, то во все моменты времени численности популяций не меняются;

б) при малом отклонении от положения равновесия численности как хищника, так и жертвы с течением времени не возвращаются к равновесным значениям (при этом из модели (1) получается стандартное уравнение колебаний; см. п. 3 § 5 гл. I);

в) если отклонение от положения равновесия велико, то поведение функций $N(t)$, $M(t)$ такое же, как в случае б).

Эти выводы означают, что численности популяций жертвы и хищника совершают периодические колебания вокруг положения равновесия. Амплитуда колебаний и их период определяются начальными значениями численностей $N(0)$, $M(0)$ (упр. 1, 2), они совершаются в противофазе: максимальному значению $N(t)$ соответствует минимальное значение $M(t)$, и наоборот (рис. 57). Колебания, сущность которых вполне понятна (и они реально наблюдаются в природе) означают возникновение в двухвидовых популяционных системах значительно более сложных процессов, чем в одновидовых системах (ср. с моделью Мальтуса и логистической моделью в гл. I).

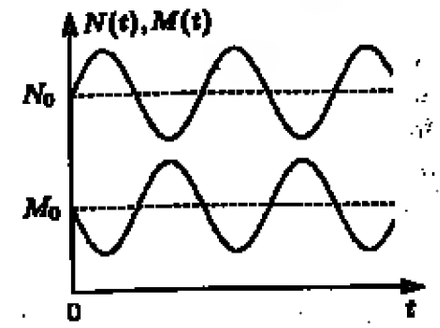


Рис. 57

Более точные математические описания двухвидовых взаимодействий учитывают неравномерность распределения численности популяций на занимаемых территориях (им соответствуют системы уравнений в частных производных), временное запаздывание между рождением особей и их зрелостью и т. д. Возникают гораздо более сложные картины взаимодействия как по времени, так и в пространстве. Например, при учете насыщения для популяции жертв в первом уравнении (1) появляется соответствующий логистический член, и оно приобретает вид

$$\frac{dN}{dt} = (\alpha - cM - aN)\,N. \tag{6}$$

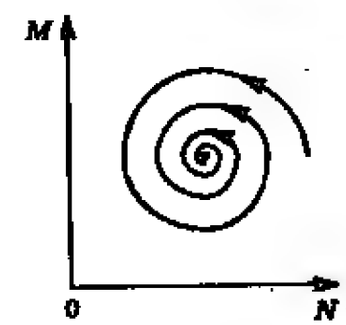


Рис. 58

В этом случае фазовые траектории имеют вид спиралей (рис. 58), сходящихся с течением времени к положению равновесия, а амплитуда колебаний уменьшается с течением времени (упр. 3).

2. Гонка вооружений между двумя странами. Предполагается, что общее количество вооружений у каждой страны изменяется со временем в зависимости от трех факторов: количества оружия у противника, износа уже существующего вооружения и степени недоверия между противниками. Темпы прироста и уменьшения вооружений пропорциональны указанным факторам, т. е.

$$\begin{aligned} \frac{dM_1}{dt} &= \alpha_1(t)\,M_2 - \beta_1(t)\,M_1 + \gamma_1(t), \\ \frac{dM_2}{dt} &= \alpha_2(t)\,M_1 - \beta_2(t)\,M_2 + \gamma_2(t). \end{aligned} \tag{7}$$

В уравнениях (7) $M_1(t) \geqslant 0$, $M_2(t) \geqslant 0$ — объемы вооружений, коэффициенты $\alpha_1(t) > 0$, $\alpha_2(t) > 0$, $\beta_1(t) > 0$, $\beta_2(t) > 0$ характеризуют скорость «старения» вооружений (аналог процесса амортизации производственных мощностей в моделях экономики), функции $\gamma_1(t) \geqslant 0$,

$\gamma_2(t) \geqslant 0$ описывают уровень взаимной настороженности (недоверия) конкурентов, который считается не зависящим от количества вооружений, а определяется другими причинами.

Модель (7) не учитывает многие важные факторы, влияющие на динамику гонки вооружений, но, тем не менее, дает возможность проанализировать ряд существенных свойств этого процесса. Анализ наиболее прост в частном случае, когда функции α_i, β_i, γ_i, $i = 1, 2$, не зависят от времени:

$$\begin{aligned} \frac{dM_1}{dt} &= \alpha_1 M_2 - \beta_1 M_1 + \gamma_1, \\ \frac{dM_2}{dt} &= \alpha_2 M_1 - \beta_2 M_2 + \gamma_2. \end{aligned} \tag{8}$$

Исследуем систему (8) в плоскости M_1, M_2 с целью определить качественное поведение функций $M_1(t)$, $M_2(t)$ во времени. Уравнения (8) имеют положение равновесия $dM_1/dt = 0$ и $dM_2/dt = 0$. Равновесные значения M_1^0, M_2^0 находятся, очевидно, из условий

$$\alpha_1 M_2 - \beta_1 M_1 + \gamma_1 = 0, \quad \alpha_2 M_1 - \beta_2 M_2 + \gamma_2 = 0$$

и равны

$$M_1^0 = \frac{\alpha_1\gamma_2 + \beta_2\gamma_1}{\beta_1\beta_2 - \alpha_1\alpha_2}, \quad M_2^0 = \frac{\alpha_2\gamma_1 + \beta_1\gamma_2}{\beta_1\beta_2 - \alpha_1\alpha_2}. \tag{9}$$

Из (9) следует первый важный вывод: для того чтобы равновесие существовало при положительных значениях величин M_1^0, M_2^0 (по своему смыслу функции $M_1(t)$, $M_2(t)$ неотрицательны), должно выполняться неравенство

$$\beta_1\beta_2 > \alpha_1\alpha_2. \tag{10}$$

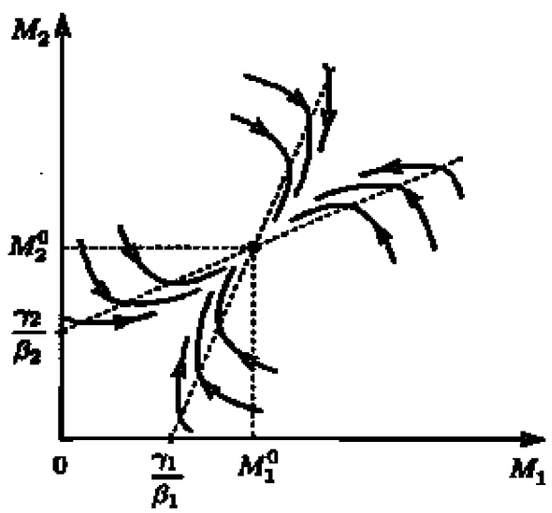


Рис. 59

Смысл условия (10) проясняется из следующих рассуждений. Пусть, например, параметры α_1, β_1 и β_2 неизменны, а параметр α_2 увеличивается. Это означает, что первая страна не меняет свою стратегию в области вооружений, а вторая наращивает вооружения при неизменном темпе износа оружия (параметр β_2). Тогда при достаточно большой величине α_2 равновесие станет заведомо невозможным, а неравенство (10) обязательно нарушится. Заметим, что, если оба параметра γ_1, γ_2, характеризующие взаимное недоверие, равны нулю, то положению равновесия отвечает отсутствие вооружений у обеих сторон.

Изучим теперь вопрос об устойчивости равновесия (9) при условии (10). В этом случае интегральные кривые уравнения (8) в плоскости M_1, M_2 имеют вид, изображенный на рис. 59. Штриховые прямые отвечают изоклинам нуля ($M_2 = \alpha_2/\beta_2\, M_1 + \gamma_2/\beta_2$) и бесконечности ($M_2 = \beta_1/\alpha_1\, M_1 - \gamma_1/\alpha_1$). Изоклина нуля имеет наклон, меньший, чем изоклина бесконечности (это следует из неравенства (10)). Сплошным линиям соответствуют интегральные кривые. Стрелками показано направление движения по интегральным кривым с течением времени. Функции $M_1(t)$ и $M_2(t)$ при возрастании t стремятся к равновесным значениям. Таким образом, равновесие устойчиво: любое отклонение от него становится ничтожно малым по прошествии достаточно большого промежутка времени (упр. 4, 5).

Из построенной модели нетрудно определить некоторые характеристики возможных поведений соперников при переходе от одного положения равновесия к другому. Пусть, например, темп наращивания вооружений в первой и второй странах изменяется на небольшую величину $d\alpha$ ($d\alpha = d\alpha_1 = d\alpha_2$). При этом объем вооружений также изменяется, причем обе стороны желают, чтобы приращения dM_1^0 и dM_2^0 были равными и интересы обеих сторон не ущемлялись. Для величин dM_1^0, dM_2^0 из (9) получаем

$$dM_1^0 = \frac{\alpha_1\alpha_2\gamma_2 + \alpha_2\beta_2\gamma_1 + \alpha_1^2\gamma_2 + \alpha_1\beta_2\gamma_1}{(\beta_1\beta_2 - \alpha_1\alpha_2)^2}\, d\alpha,$$

$$dM_2^0 = \frac{\alpha_1\alpha_2\gamma_1 + \alpha_1\beta_1\gamma_2 + \alpha_2^2\gamma_1 + \alpha_2\beta_1\gamma_2}{(\beta_1\beta_2 - \alpha_1\alpha_2)^2}\, d\alpha.$$

Предположим для простоты, что настороженности (недоверия) партнеров равные ($\gamma_1 = \gamma_2$). Тогда из равенства $dM_1^0 = dM_2^0$ получаем условие паритета сторон при небольшом изменении равновесия

$$\alpha_1(\alpha_1 + \beta_2 - \beta_1) = \alpha_2(\alpha_2 + \beta_1 - \beta_2),$$

которое может быть положено в основу соответствующих договоренностей между странами, если известны величины α_1, α_2, β_1, β_2. Так, например, пусть $\alpha_2 = \sigma\alpha_1$, $\sigma > 0$. В этом случае из предыдущего равенства получаем

$$\alpha_1(1 - \sigma) = \beta_1 - \beta_2. \tag{11}$$

При $\sigma < 1$ (темп прироста вооружений у второй стороны меньше, чем у первой) для сохранения паритета необходимо $\beta_2 < \beta_1$, т. е. у второй стороны (в соответствии с формулой (11)) темп амортизации вооружений должен быть меньше. При противоположном неравенстве $\sigma > 1$ имеем, естественно, обратное соотношение между скоростями амортизации.

3. Боевые действия двух армий. В противоборстве могут принимать участие как регулярные армии, так и партизанские соединения. Главной характеристикой соперников в рассматриваемых моделях являются численности сторон $N_1(t) \geqslant 0$ и $N_2(t) \geqslant 0$. Если в какой-то момент времени одна из численностей обращается в нуль, то данная

сторона считается потерпевшей поражение (притом, что в этот момент численность другой стороны положительна).

В случае действий между регулярными частями динамика их численности определяется тремя факторами:

1) скоростью уменьшения состава из-за причин, непосредственно не связанных с боевыми действиями (болезни, травмы, дезертирство);

2) темпом потерь, обусловленных боевыми действиями противоборствующей стороны (которые в свою очередь определяются качеством ее стратегии и тактики, уровнем морального духа и профессионализмом бойцов, вооружениями и т. д.);

3) скоростью поступления подкреплений, которая считается некоторой заданной функцией времени.

При этих предположениях для $N_1(t)$, $N_2(t)$ получаем систему уравнений

$$\begin{aligned} \frac{dN_1}{dt} &= -\alpha_1(t)\,N_1 - \beta_2(t)\,N_2 + \gamma_1(t), \\ \frac{dN_2}{dt} &= -\alpha_2(t)\,N_2 - \beta_1(t)\,N_1 + \gamma_2(t), \end{aligned} \tag{12}$$

из которой при заданных функциях α_i, β_i, γ_i, $i = 1, 2$, и начальных значениях $N_1(0)$, $N_2(0)$ однозначно определяется решение в любой момент времени $t > 0$. В (12) коэффициенты $\alpha_{1,2}(t) \geqslant 0$ характеризуют скорости потерь в силу обычных (не боевых) причин, $\beta_{1,2}(t) \geqslant 0$ — темпы потерь из-за действий соперника, $\gamma_{1,2} \geqslant 0$ — скорости поступления подкреплений.

Война между регулярными и партизанскими частями описывается другой моделью. Главное отличие в том, что нерегулярные соединения в сравнении с армейскими менее уязвимы, так как действуют скрытно, зачастую оставаясь невидимыми для соперника, вынужденного действовать неизбирательно, по площадям, занимаемым партизанами. Поэтому считается, что темп потерь партизан, проводящих свои операции в разных местах на некоторой известной территории, пропорционален не только численности армейских соединений $N_1(t)$, но и численности самих партизан $N_2(t)$, т. е. определяется членом вида $\beta_1(t)\,N_1 N_2$. В результате модель становится нелинейной:

$$\begin{aligned} \frac{dN_1}{dt} &= -\alpha_1(t)\,N_1 - \beta_2(t)\,N_2 + \gamma_1(t), \\ \frac{dN_2}{dt} &= -\alpha_2(t)\,N_2 - \beta_1(t)\,N_1 N_2 + \gamma_2(t). \end{aligned} \tag{13}$$

В (13) все величины имеют тот же смысл, что и в (12). Изучим модели (12), (13) (*модели Ланчестера*) в частном случае: $\gamma_1 = \gamma_2 = 0$ (стороны не получают подкреплений и как бы предоставлены самим себе); $\alpha_1 = \text{const}$, $\alpha_2 = \text{const}$; $\beta_1 = \text{const}$, $\beta_2 = \text{const}$ (последнее означает, в частности, что у противников всегда найдется достаточное количество вооружений, которое может использоваться годными к несению службы бойцами).

Модель (12) становится автономной и принимает вид

$$\begin{aligned} \frac{dN_1}{dt} &= -\alpha_1 N_1 - \beta_2 N_2, \\ \frac{dN_2}{dt} &= -\alpha_2 N_2 - \beta_1 N_1. \end{aligned} \tag{14}$$

Из уравнений (14) видно, что в данном случае численности сторон с течением времени могут только убывать. Каков временной характер этого процесса и какая сторона потерпит поражение? Чтобы выяснить этот вопрос, введем еще одно упрощение (вполне оправданное для краткосрочных кампаний): положим $\alpha_1 = \alpha_2 = 0$. Другими словами, потери сторон определяются лишь действиями противника. Система (14) упрощается:

$$\frac{dN_1}{dt} = -\beta_2 N_2, \quad \frac{dN_2}{dt} = -\beta_1 N_1, \tag{15}$$

и легко находится ее интеграл

$$\beta_1 N_1^2(t) - \beta_2 N_2^2(t) = \beta_1 N_1^2(0) - \beta_2 N_2^2(0) = C. \tag{16}$$

Из (16) однозначно определяется победитель (рис. 60). При $C > 0$ побеждает первая армия, при $C < 0$ — вторая, в случае $C = 0$ стороны уничтожают друг друга одновременно, и победителя нет. Смысл этих результатов вполне ясен из вида константы в (16). Для победы важна не только численность сторон в начале боевых действий ($N_1(0)$, $N_2(0)$), но и их выучка, качество их вооружений и т. д. (т. е. коэффициенты β_1, β_2). Так, если $C > 0$, то из (16) следует

$$\beta_1 N_1^2(0) > \beta_2 N_2^2(0);$$

и для достижения победы второй стороне следует либо увеличить начальную численность, либо улучшить качество боевых действий, либо то и другое одновременно. Заметим, что эффект от увеличения коэффициента β_2 меньше, чем от такого же увеличения числа $N_2(0)$, которое входит в последнее неравенство (так называемый *квадратичный закон боевых действий*) во второй степени.

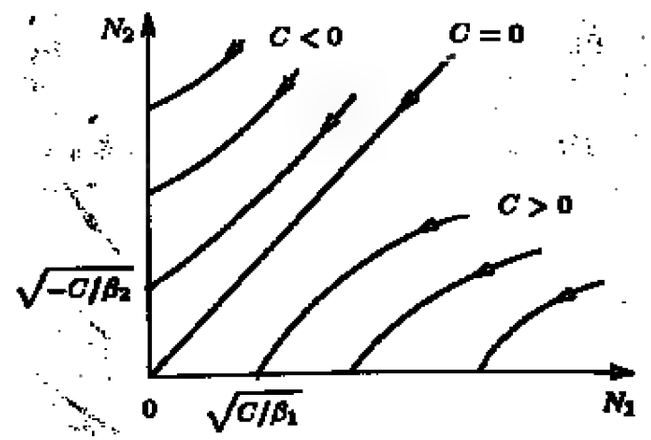


Рис. 60

Дифференцируя первое из уравнений (15) и принимая во внимание второе, получаем уравнение для $N_1(t)$

$$\frac{d^2 N_1}{dt^2} = \beta_1 \beta_2 N_1. \tag{17}$$

Из (17) с учетом начальных условий $N_1(t=0)=N_1(0)$ и $dN_1/dt\,(t=0)=-\beta_2 N_2(0)$ находим численность первой армии как функцию времени:

$$N_1(t)=N_1(0)\operatorname{ch}\sqrt{\beta_1\beta_2}\,t-\sqrt{(\beta_2/\beta_1)\,N_2(0)}\operatorname{sh}\sqrt{\beta_1\beta_2}\,t, \qquad (18)$$

зная которую, нетрудно найти и функцию $N_2(t)$ (упр. 7).

Рассмотрим теперь действия регулярной армии против партизан в тех же упрощающих предположениях, что и в предыдущем случае. Модель (13) приобретает вид

$$\frac{dN_1}{dt}=-\beta_2 N_2,\quad \frac{dN_2}{dt}=-\beta_1 N_1 N_2. \qquad (19)$$

Численности сторон, как и прежде, убывают со временем, но по другому закону. Умножим первое уравнение (19) на $\beta_1 N_1$, второе — на β_2 и вычтем второе получившееся уравнение из первого. В итоге придем к уравнению

$$\frac{d}{dt}\left[\frac{\beta_1}{2}N_1^2(t)-\beta_2 N_2(t)\right]=0,$$

имеющему интеграл

$$\frac{\beta_1}{2}N_1^2(t)-\beta_2 N_2(t)=\frac{\beta_1}{2}N_1^2(0)-\beta_2 N_2(0)=C_1. \qquad (20)$$

Изучим фазовые траектории системы (19) с помощью интеграла (20). Из рис. 61. видно, что при $C_1>0$ побеждает армия, при $C_1<$

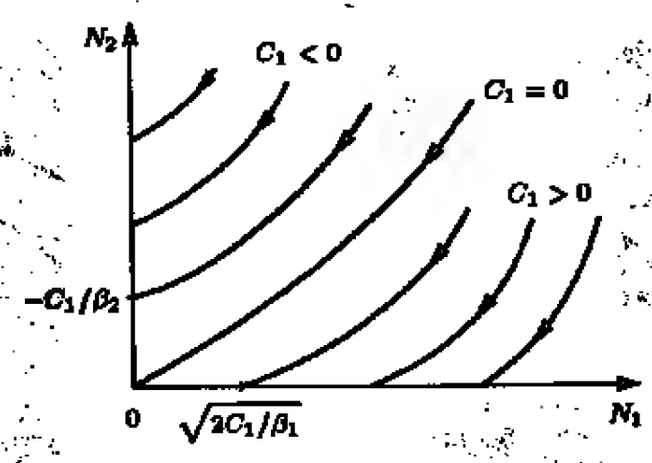


Рис. 61

<0 — партизаны, а при $C_1=0$ победителя нет. Так же, как и при действии регулярных частей, победа обеспечивается не только начальной численностью, но и боевой выучкой и качеством оружия. Пусть, например, $C_1>0$, т. е.

$$\frac{\beta_1}{2}N_1^2(0)>\beta_2 N_2(0). \qquad (21)$$

Тогда партизанам необходимо обеспечить увеличение коэффициента β_2 и повысить свою начальную численность $N_2(0)$ на соответствующую величину, иначе они обречены на поражение. Причем это увеличение с ростом величины $N_1(0)$ должно расти не линейно, а пропорционально второй степени $N_1(0)$ (*параболический закон боевых действий*). Можно сказать, что в некотором смысле регулярные соединения находятся в более выгодном положении, поскольку неравенство (21) для них достигается при меньшем росте начальной численности, чем противоположное (21) неравенство для численности партизан.

Поведение функций $N_1(t)$, $N_2(t)$ во времени находится из (19) с использованием интеграла (20). Так, для $N_1(t)$ имеем

$$\frac{dN_1}{dt}=C_1-\frac{\beta_1}{2}N_1^2,$$

что эквивалентно уравнению

$$\frac{dN_1}{C_1-\beta_1 N_1^2/2}=dt. \qquad (22)$$

Интегрируя (22), нетрудно найти $N_1(t)$ и затем $N_2(t)$ как неявные функции времени (упр. 7).

В заключение отметим, что рассмотренные здесь простейшие модели соперничества соответствуют системам обыкновенных дифференциальных уравнений второго порядка (в общем случае неавтономных и нелинейных), широко распространенным при описании многих естественнонаучных объектов. Это закономерно, так как использованные при построении подходы (насыщение, пропорциональность темпов роста величины значению этой величины и др.) аналогичны подходам, применяемым в механике, физике, химии.

УПРАЖНЕНИЯ

1. Покажите, линеаризуя систему (1) в окрестности равновесия (3), что соответствующая ему особая точка является центром.

2. Вычислите период колебаний в системе «хищник—жертва» в зависимости от ее характеристик (α, β, c, d) и начального состояния.

3. Убедитесь, линеаризуя систему уравнений, состоящую из уравнения (6) и второго уравнения (1), в том, что для нее точка равновесия — устойчивый фокус.

4. Покажите, что точка равновесия (9) — устойчивый узел.

5. Получите оценку для времени прихода системы (8) в равновесие с точностью до 1% в зависимости от ее характеристик и величины начального отклонения.

6. Установите характер кривизны фазовых траекторий на рис. 60.

7. Найдите функцию $N_2(t)$, используя либо уравнение (15), либо интеграл (16), и сравните ее с функцией $N_1(t)$ из (18). Убедитесь в том, что выводы, сделанные на основе анализа интеграла (16), правильны (т. е. что при $C>0$ побеждает первая сторона, при $C<0$ — вторая).

8. Найдите решение уравнения (22) и определите момент, когда функция $N_1(t)$ обращается в нуль (при этом, естественно, $C_1<0$). Установите характер кривизны фазовых траекторий на рис. 61.

§ 4. Динамика распределения власти в иерархии

Построим и изучим макромодель, описывающую некоторые ключевые взаимодействия в системе «государственная власть—гражданское общество». Рассмотрим влияние реакции гражданского общества и других характеристик системы на динамику распределения власти внутри иерархии, а также некоторые свойства этих распределений.

1. Общая постановка задачи и терминологии. Изучение структур власти — одна из главных задач наук об обществе, прежде всего политологии. Понятие «власть» таинственно, многозначно, с трудом поддается формализации и количественному измерению. Естественно поэтому, что математические модели политологии носят в основном описательный, феноменологический характер и применяются к относительно узкому кругу проблем мониторинга: статистическая обработка электоральных ожиданий и результатов выборов, определение рейтингов различных политических сил, прогнозы грядущих парламентских голосований на основе уже состоявшихся и т. д.

Построение математических моделей общей политологии разумно начинать с изучения именно государственных иерархий как обладающих властью на официальных, т. е. легко формализуемых основаниях. В этом состоит важное отличие государственной власти от власти независимых средств массовой информации (СМИ), интеллектуальных и нравственных авторитетов и других видов власти.

Под *иерархией*, или *иерархической структурой*, понимается упорядоченная по старшинству совокупность институтов (инстанций, должностей, постов, чинов и т. п.), наделенных властными полномочиями от имени государства (т. е. по Конституции, законам, уставам, постановлениям, правилам, инструкциям и т. п.).

Имеются в виду не только общегосударственные учреждения, например, федеральные министерства, но также межрегиональные, региональные и местные органы, официально имеющие соответствующую власть. Слово «иерархическая» подчеркивает, что внутри структуры заранее и четко определен порядок подчиненности. Каждое ее звено (кроме высшего) имеет старшие, «приказывающие» инстанции и (за исключением низшего звена) младшие инстанции, выполняющие «приказы», исходящие как от данной инстанции, так и от других старших звеньев. Разумеется, приказы идут только от старших к младшим.

Гражданское общество — это часть общества, непосредственно не обладающая государственной властью. К нему относятся граждане (в том числе государственные чиновники, действующие вне рамок своих служебных обязанностей) и различные их объединения (политические, культурные, профессиональные), семьи и частные предприятия и т. д. Очевидно, что члены гражданского общества не могут «приказывать» от имени государства ни друг другу, ни каким-либо звеньям властной структуры. Например, негосударственная корпорация, какой бы крупной она ни была, не вправе официально принуждать никакое частное лицо к тому или иному поведению. В то же время любая инстанция имеет такую возможность по отношению к определенной части гражданского общества, а некоторые — по отношению ко всему обществу. Заметим, что одно и то же «физическое лицо» может одновременно как принадлежать ко властной структуре, так и быть членом гражданского общества.

Реакция общества — ответ (положительный, отрицательный или безразличный) гражданского общества на действия того или иного института власти (с помощью выборов, референдумов, плебисцитов, через СМИ, опросы общественного мнения, митинги, забастовки т. д.).

В модель вводится также понятие «реакция иерархии», смысл которого, как и смысл термина «распределение власти» и другие терминологические уточнения, разъясняется ниже.

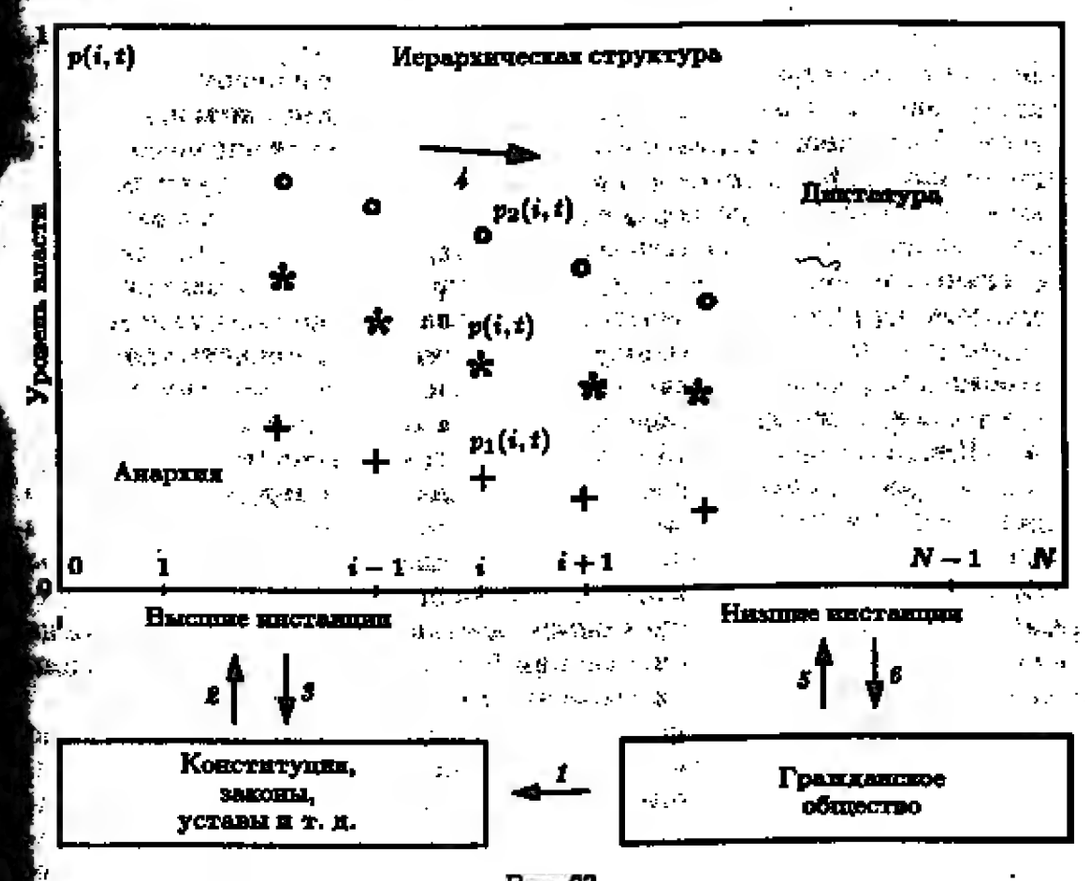


Рис. 62

Иерархическая структура состоит из $N+1$ инстанции (рис. 62), каждой из них приписан номер i $(0 \leqslant i \leqslant N)$. Высшая инстанция имеет номер 0, низшая — номер N. Стрелка 4 обозначает направление подчиненности в структуре «сверху вниз» (или «слева направо»), т. е. направление движения властных распоряжений (приказов), передаваемых по иерархической лестнице.

Конечно же, для любой конкретной и достаточно протяженной иерархической структуры (тем более в масштабах государства, пусть небольшого) определить расположение всех составляющих ее инстанций в порядке подчиненности — очень сложная и трудоемкая задача. В том числе и потому, что властные органы имеют в реальности «дре-

весное» строение, могут принадлежать разным ведомствам, быть как персонифицированными, так и коллегиальными и т. д. Понятие «подчиненность» имеет поэтому условный осредненный смысл.

Однако для построения математической модели и анализа фундаментальных свойств властных структур достаточно того, что эта задача может быть решена в принципе, впрочем, на основе некоторых количественных критериев. Для прояснения этого вопроса необходимо уточнить различные понимания слова *власть*.

Оно часто употребляется в смысле *органы власти* (Верховный Суд, городская дума, местное отделение милиции т. п.). Равноценной заменой данного понятия служат термины «инстанции», «иерархическая структура», «иерархия».

Другое важное значение этого слова описывается термином *властные полномочия*. Считается, что властные полномочия любой инстанции могут быть, вообще говоря, определены с помощью некоторого совокупного количественного критерия, включающего в себя номинальное (формальное) положение инстанции в структуре, объем находящихся в ее распоряжении людских, финансовых, материальных, информационных, интеллектуальных, законодательных, распорядительных и иных видов ресурсов, размер и местоположение контролируемой территории, престижность в глазах общественного мнения и экспертов и т. д. В указанном понимании слово «власть» означает возможный уровень (степень, силу) влияния данного властного института на поведение других инстанций и на жизнь гражданского общества. Инстанции, имеющей большие властные полномочия, отводится, естественно, более высокое место (меньший номер) в иерархии по сравнению с инстанцией, имеющей меньшие полномочия. При этом для рассмотрения общих вопросов нет необходимости вводить какие-либо абсолютные единицы измерения власти; достаточно принять властные полномочия высшей инстанции за единицу (или за 100%), тогда полномочия любой другой инстанции будут выражаться в долях (или процентах) по отношению к высшему институту.

Дальнейшее уточнение термина «властные полномочия» связано с понятиями *максимальные полномочия* и *минимальные полномочия*. Первое описывает те действия, которые орган власти может выполнить максимально в некоторой ситуации в соответствии с законодательством. Например, губернатор имеет право, при известных обстоятельствах, объявить чрезвычайное положение на контролируемой им территории, но никак не правомочен решать вопросы войны и мира. Минимальные полномочия описывают действия, которые власть обязана предпринимать всегда (Президент обязан представлять ежегодное бюджетное послание в Конгресс). Оба эти понятия хорошо иллюстрируются, например, статьями Уголовного кодекса, определяющими за один и тот же вид преступления максимальные и минимальные сроки наказания. Другими словами, максимальные и минимальные полномочия дают законные верхние и нижние границы власти каждой инстанции.

Заметим, что границы власти могут, как известно из практики многих государств, определяться не только законодательством, но также традициями и системой прецедентов. Однако для краткости всюду дальше речь будет идти о законодательно установленных полномочиях.

В математической модели максимальные властные полномочия задаются некоторой положительной функцией $p_2(i,t)$, $0 \leqslant i \leqslant N$ (кружочки на рис. 62), монотонно убывающей с ростом номера i, т. е. в любой момент времени t справедливо $p_2(i+1,t) < p_2(i,t) < p_2(i-1,t)$, $1 \leqslant i \leqslant N-1$. Минимальные полномочия задаются положительной функцией $p_1(i,t)$, $0 \leqslant i \leqslant N$ (крестики на рис. 62), также монотонно убывающей с ростом i, т. е. $p_1(i+1,t) < p_1(i,t) < p_1(i-1,t)$, $1 \leqslant i \leqslant N-1$ (разумеется, всегда $p_1(i,t) < p_2(i,t)$, $0 \leqslant i \leqslant N$). Обе функции, вообще говоря, зависят от времени t, поскольку со временем могут изменяться законодательство, территориальное деление и т. п.

Чисто умозрительные ситуации, когда в законодательстве записано «власть может делать все» или «власть не обязана делать что-либо», естественно, не рассматриваются, т. е. границы власти (хотя бы формально) всегда определены, и вместе с ними определены функции p_1, p_2.

Подчеркнем еще раз: для рассматриваемой таким образом (в агрегированной обобщенной манере) иерархической структуры положение любой инстанции определяется не только ее номинальным местом, но и всеми относящимися к делу (и зачастую более весомыми) факторами. Таким образом, номер инстанции, отвечающий некоторому «среднему чиновнику» из властного древа, не формальная координата, а координата «по существу». Приведенным выше рассуждениям о властных полномочиях и о способе расстановки инстанций в порядке «подчиненности» нетрудно придать более подробную и строгую математическую форму (микроописание), однако для изучения принципиальных свойств системы «власть—общество» это не имеет решающего значения.

Наконец, еще одно используемое в модели понимание слова «власть» связано с фактически осуществляемым в данный момент данной инстанцией *уровнем властного влияния* (или *величиной власти*). Действительно, властные полномочия определяют лишь верхние и нижние законные границы уровня или величины власти (в этом смысле можно понимать известное выражение «объем властных полномочий»). Эти границы, вообще говоря, достигаются не всегда и не всюду. Пусть, например, по некоторой статье Уголовного кодекса, предусматривающей от трех до пяти лет лишения свободы, суды за некоторое время вынесли достаточно много приговоров со средним сроком четыре года. Тогда реализованная ими по данной статье власть составила 80% от их максимальных и 133% от их минимальных полномочий.

В математической модели фактически достигаемой власти отвечает неотрицательная функция $p(i,t)$, $0 \leqslant t \leqslant N$, зависящая от «пространственной» координаты i и времени t (звездочки на рис. 62). Если для каких-то значений i, t выполняется $p(i,t) > p_2(i,t)$ (или $p(i,t) < p_1(i,t)$), то естественно говорить о «выходе» власти за рамки полномочий, или о *превышении* (*принижении*) *власти*.

Принципиальная разница между функциями $p_1(i,t)$, $p_2(i,t)$ и функцией $p(i,t)$ состоит в том, что, в отличие от известных, заданных (пусть и в самом общем виде) властных полномочий $p_1(i,t)$, $p_2(i,t)$, функция $p(i,t)$ — неизвестная, искомая величина, описывающая *текущее распределение власти в иерархической структуре*. Построение соответствующей математической модели и изучение пространственно-временной динамики именно для функции $p(i,t)$ (распределения власти) в зависимости от всех факторов, присутствующих в изучаемой системе, приводится ниже.

Заметим, что понятие *реальное распределение власти* означает введение в модель следующего предположения 1:

все партнеры в системе «власть—общество» законопослушны: законы соблюдаются, налоги платятся, приказы выполняются (в противном случае функция $p(i,t)$ — величина власти — становится весьма неопределенной или вовсе теряет смысл).

Такой подход «от простого к сложному», т. е. построение модели для сильно идеализированной ситуации и дальнейшее ее усовершенствование, является типичным (а зачастую и единственно возможным) при математическом моделировании сложных объектов.

Приведем общее описание взаимодействий в системе «власть—общество». Оно состоит в следующем:

а) гражданское общество непосредственно или через своих представителей принимает (дополняет, изменяет) Конституцию (стрелка *1* на рис. 62). Оно, таким образом, выступает как источник (заказчик, хозяин) власти для иерархической структуры, взаимодействуя с нею с учетом существующего законодательства;

б) иерархическая структура существует не сама по себе, а как «открытая» система, взаимодействующая с Конституцией и гражданским обществом. Конституция (в широком понимании, включающем законы, уставы и т. д.) служит для иерархии своеобразным резервуаром, из которого ее звенья по мере необходимости могут либо черпать дополнительные «порции» власти (стрелка *2*), либо возвращать обратно ее избыточные «доли» (стрелка *3*).

Таким образом, между иерархией и Конституцией осуществляется как бы «обмен властью» и, в неявном виде, как бы обмен властью («свободой») между иерархией и гражданским обществом — учредителем Конституции;

в) внутри самой иерархической структуры происходит перераспределение текущей власти между составляющими ее инстанциями (стрелка *4*) в соответствии с принятыми в иерархии механизмами передачи властных распоряжений;

г) по отношению к гражданскому обществу властные структуры выступают как запрещающие (стрелка *5*) или разрешающие (стрелка *6*) институты, вводящие и отменяющие те или иные ужесточения или послабления (типичный пример — ежегодный призыв в армию и демобилизация, когда одна часть населения начинает нести дополнительную повинность, а другая от нее освобождается).

Один из ключевых вопросов для описания взаимодействия в системе «иерархия—общество» — определение «величины» обмена властью

между иерархической структурой и Конституцией (и, опосредованно, в итоге между иерархией и обществом). Вводится следующее предположение 2:

знак и величина обмена властью между иерархической структурой и Конституцией определяются реакцией системы.

Под *реакцией системы* подразумевается суммарная реакция обоих партнеров (иерархии и общества) на текущее распределение власти $p(i,t)$ в структуре. Например, применительно к *реакции общества* имеется в виду, что если в данный момент времени оно выражает свое отрицание тех или иных действий данного звена иерархии (сопротивление), то это побуждает инстанцию уменьшить осуществляемую ею власть, как бы отложив про запас некоторое ее «количество» в конституционный резервуар (и уменьшить или вовсе избежать соответствующих отрицательных, с точки зрения общества, последствий). Подходящий пример — требования смягчить те или иные виды налогов. Положительная же реакция общества (т. е. его поддержка) побуждает инстанцию увеличить уровень реализуемой ею власти, взяв необходимые «ресурсы» из законодательства (пример — требования усилить борьбу с преступностью). Качественный характер реакции общества связывается с доминирующим в нем типом сознания (правовое, анархическое, авторитарное, смешанное).

В рассматриваемой модели реакция общества описывается некоторой задаваемой функцией $F_S(i,t,p,p_1,p_2)$, зависящей, вообще говоря, от всех введенных ранее величин: номера инстанции i; времени t; уровня власти, осуществляемой инстанцией $p(i,t)$; властных полномочий $p_1(i,t)$, $p_2(i,t)$. Это дает возможность достаточно полно и точно отразить структурированное и меняющееся во времени отношение общества к иерархии. Например, если для всех мыслимых значений своих аргументов, т. е. в любой ситуации имеем $F_S < 0$ ($F_S > 0$), то в обществе, очевидно, преобладает анархическое (тоталитарное) сознание. Реакция общества представляется величиной, вполне наблюдаемой и измеряемой, и служит основной его поведенческой характеристикой в некотором осредненном по всем элементам гражданского общества смысле. При этом подразумевается, что реакция выражена вовремя, верно истолкована инстанцией и учтена ею в своей деятельности.

Аналогичным образом в модель вводится одна из поведенческих характеристик властной структуры — *реакция иерархии*. Это понятие характеризует стремление звеньев иерархии повысить или понизить уровень власти, реализуемый ими в данный момент времени. Ей отвечает задаваемая функция $F_H(i,t,p,p_1,p_2)$, зависящая от тех же аргументов, что и F_S, и имеющая тот же смысл, но уже применительно к иерархии (ее можно назвать *степенью «властолюбия» инстанций*).

В дальнейшем для простоты будем считать $F_H \equiv 0$, т. е. иерархии безразличен уровень осуществляемой ею власти, и реакция системы $F = F_S + F_H$ определяется лишь «гражданской» компонентой. Данное упрощение не меняет чисто математических свойств модели, поскольку функции F_S, F_H, F зависят от одних и тех же аргументов. Однако при интерпретации результатов, разумеется, необходимо учитывать роль каждой составляющей F_S, F_H в суммарной реакции F.

В приведенном общем описании система «власть—общество» предстает как замкнутый самосогласованный и самоорганизующийся объект с различными прямыми и обратными связями. Уровень власти, осуществляемой любой инстанцией иерархии в любой момент времени, отнюдь не произволен, а является результатом системного взаимодействия всех компонент объекта: идущих по иерархической лестнице властных распоряжений, реакции общества, действующего законодательства, начального состояния системы и т. д. Для получения на основе этой общей схемы математической модели рассмотрим подробнее собственно властную структуру.

2. Механизмы перераспределения власти внутри иерархической структуры. Внутри иерархии любая инстанция принимает к исполнению какие-то властные распоряжения, идущие от старших звеньев, и в свою очередь передает какие-то приказы младшим звеньям. При этом происходит некоторое перераспределение власти между ступенями иерархии (напомним, что речь идет не о властных полномочиях, а о реально осуществляемом, текущем уровне власти — величине $p(i,t)$). Вводится основной поведенческий постулат:

в иерархии власть может передаваться только от инстанций с большей текущей властью к инстанциям с меньшей текущей властью (причем скорость передачи тем больше, чем больше разница между значениями текущей власти в инстанциях).

Рассматриваются два главных механизма, ответственных за перераспределение власти внутри иерархии.

а) *Близкодействие.* Условно говоря, этот механизм можно назвать *передачей власти по команде*, когда начальник отдает властное распоряжение ближайшим подчиненным, которые также взаимодействуют только со своими непосредственными подчиненными. Данный механизм соответствует хорошо известной бюрократической процедуре (слово «бюрократический» и все другие понятия употребляются без каких-либо эмоциональных оттенков — как рабочий термин).

Пусть i-я инстанция передала $(i+1)$-й инстанции некий приказ (например, поручила подготовить проект годового отчета о финансовой деятельности одного из подведомственных учреждений). Что произошло в таком элементарном акте взаимодействия между соседями по иерархии? Вместе с приказом подчиненный получил некоторую (пусть небольшую и пусть временно) порцию власти, дополнительную к тому уровню власти, который он реально занимал (например, имея приказ, он в большей степени, чем раньше, изучает финансовые документы). С другой стороны, i-я инстанция утратила некоторую часть от своей текущей власти, переложив контроль за данным участком работы на $(i+1)$-е звено.

Совокупность передаваемых распоряжений формирует своего рода поток власти, идущей от i-й к $(i+1)$-й инстанции. Определим *поток власти* $W(i,t)$ как количество власти, полученное в единицу времени $(i+1)$-й ступенью от i-й ступени.

В соответствии с постулатом величина $W(i,t)$ положительна при $p(i,t) > p(i+1,t)$, отрицательна при $p(i,t) < p(i+1,t)$ и равна нулю

при $p(i,t) = p(i+1,t)$. Рассматривается весьма общее выражение для $W(i,t)$:

$$W(i,t) = -\varkappa[p(i,t) - p(i+1,t), p(i,t), p(i+1,t), i, i+1, t] \times \\ \times [p(i+1,t) - p(i,t)], \qquad (1)$$

где функция $\varkappa$ положительна при всех значениях аргументов, от которых она зависит.

Формула (1) — конкретизация постулата применительно к рассматриваемому механизму близкодействия. Положительность функции $\varkappa$ обеспечивает нужный знак величины $W(i,t)$. В частности, если в какой-то момент времени t оказывается, что $p(i+1,t) > p(i,t)$, т. е. младшая инстанция имеет власть, большую, чем старшая, то поток власти автоматически «срабатывает» в сторону уменьшения разницы $p(i,t) - p(i+1,t)$, способствуя преодолению этой «ненормальной» ситуации (свидетельствующей о неблагополучии в данных звеньях иерархии). Поток власти растет с ростом разницы $p(i,t) - p(i+1,t)$, что соответствует второй части постулата (для этого требуется наложить соответствующее ограничение на $\varkappa$ как функцию от $p(i,t) - p(i+1,t)$). Это означает, в частности, что начальник (при прочих равных условиях) дает тем больше поручений подчиненному, чем меньше тот в данный момент времени загружен работой.

Общность выражения (1) в том, что для потока власти предусмотрена, помимо зависимости от разницы $p(i,t) - p(i+1,t)$, также зависимость от всех введенных ранее величин: значений искомой функции $p(i+1,t), p(i,t)$; координат $i, i+1$; времени t. Зависимость от t и $i, i+1$ означает также неявную зависимость от заданных властных полномочий p_1, p_2 и реакции общества F (впрочем, эту зависимость нетрудно ввести и явным образом). Общность заключается также и в том, что на все эти зависимости не накладывается никаких существенных ограничений, кроме требования, чтобы при $p(i+1,t) = p(i,t)$ выполнялось $W(i,t) = 0$. Заметим, однако, что по формуле (1) поток власти определяется только величинами, относящимися к непосредственно взаимодействующим (в данном случае соседним) инстанциям.

Функция $\varkappa$ имеет определенный бюрократический смысл, описывая свойства властной «среды», некоторые аспекты взаимоотношений внутри иерархии. Пусть, к примеру, поток власти $W(i,t)$ и величина $p(i,t)$ фиксированы, а $\varkappa = \varkappa_0 = \text{const}$, т. е. функция $\varkappa$ не зависит от времени и других величин. Тогда по формуле (1) получаем, что с возрастанием $\varkappa_0$ разница $p(i+1,t) - p(i,t)$ как абсолютно, так и относительно уменьшается. Можно сказать, что с ростом $\varkappa_0$ уменьшается *«отстаственность»* старшей инстанции, которая все более терпимо относится к выравниванию своей власти с властью младшего партнера. С уменьшением величины $\varkappa_0$ ситуация, естественно, обратная.

б) *Дальнодействие.* Образное описание этого механизма может быть дано известным выражением *«команда через голову»*. Это означает, что i-я инстанция отдает властные распоряжения звеньям с далекими номерами, минуя ближайших подчиненных. К такого рода действиям можно отнести, например, распоряжение о переходе вооруженных сил на летнюю форму одежды, для реализации которого

фактически не требуется процедура передачи по команде, достаточно опубликовать соответствующий приказ для сведения (и исполнения) сразу всех военнослужащих.

Образующийся благодаря механизму дальнодействия поток власти (смысл тот же, что и в случае а)), получаемый $(i+1)$-й инстанцией от j-го звена, может быть выражен достаточно общей формулой

$$V(i+1,j,t) = \chi(p(i+1,t), p(j,t), i+1, j, t) \cdot [p(j,t) - p(i+1,t)], \quad (2)$$

где функция $\chi \geqslant 0$ в согласии с постулатом. По своему смыслу поведенческая характеристика χ близка к функции $\varkappa$, но, во-первых, ею описывается взаимодействие не соседних, а удаленных друг от друга инстанций, т. е. $j \neq i+2, i$; во-вторых, величина χ может обращаться в нуль, т. е. предусмотрена возможность отсутствия команд через голову между какими-либо ступенями иерархии.

Так же, как и функция $\varkappa$, она зависит только от величин, относящихся к непосредственно взаимодействующим звеньям. Эти зависимости априори не подвергаются каким-либо ограничениям (требуется только, чтобы $V(i+1,j,t) = 0$ при $p(i+1,t) = p(j,t)$) и могут быть дополнены явной зависимостью от других введенных ранее величин.

Подчеркнем еще раз, что оба рассмотренных механизма представляют собой усреднение, агрегирование детального микроописания взаимодействий внутри иерархической структуры, имеющей, конечно же, не простую конфигурацию цепочки, а скорее топологию большого количества деревьев со взаимопересекающимися кронами.

Сделаем еще одно важное разъяснение. В механизмах а) и б) властные распоряжения, передаваемые от i-й (j-й) к $(i+1)$-й инстанции, в нормальной ситуации (т. е. при $p(i,t) > p(i+1,t)$ или $p(j,t) > p(i+1,t)$, $j < i$) несут с собой, согласно постулату и его конкретизациям (1), (2), некоторую положительную порцию власти для $(i+1)$-й инстанции. Если же имеет место уже упоминавшийся ненормальный случай, когда $p(i,t) < p(i+1,t)$ или $p(j,t) < p(i+1,t)$, $j < i$, то по постулату некоторая доля власти переходит уже от младшего к старшему звену. Это предположение естественно и отвечает сути иерархической структуры (если, конечно, конкретный начальник не заблуждается относительно реального уровня своей власти по сравнению с реальной властью подчиненного).

В то же время постулат, описывая ситуацию «младший не может получить порцию власти от старшего, если тот в данный момент имеет меньшую (или равную) реальную власть», не описывает, казалось бы, возможную ситуацию, когда «старший имеет власть, большую, чем младший, но вдобавок еще и отбирает у младшего часть его власти» (это противоречило бы постулату).

Рассмотрим эту последнюю ситуацию более подробно, чтобы убедиться, что в точной, правильной интерпретации она фактически нереализуема (по крайней мере, в некотором «среднем» смысле).

Действительно, отобрав часть власти у кого-либо из своих подчиненных (например, отозвав некоторое поручение), начальник рано или поздно все равно «спустит вниз» это поручение в виде какого-либо приказа, но, быть может, другими путями. Ибо в любой иерархии никакая

инстанция не способна взять на себя непосредственную реализацию заметной доли поручений, с необходимостью выполняемых даже малой частью подчиненных ей звеньев (имеется в виду достаточно протяженная иерархическая структура). С другой стороны, дело в итоге должно быть сделано, и приказ, а с ним и порция власти, будут так или иначе переданы младшим. Таким образом, введенный постулат естествен и обоснован хотя бы «в среднем».

Иное дело — более конкретные взаимоотношения между звеньями иерархии, внешне похожие на обсужденную выше ситуацию. Они могут быть описаны определенными ранее функциями. Пусть, например, в условной паре «начальник—заместители начальника» первый по каким-либо причинам не удовлетворен работой одного из своих заместителей. Старший может, например, предпринять следующие действия:

1) сместить «физическое лицо», заменив его более подходящей фигурой;

2) направить поток властных распоряжений через других заместителей.

В этих двух случаях в иерархии в целом и в отвечающей ей модели фактически не меняется ничего, кроме персон. В частности, инстанция «заместители» как целое получает тот же поток власти.

В распоряжении начальника могут иметься и такие способы:

3) усилить роль команд через голову данного заместителя;

4) изменить его властные полномочия;

5) упразднить данную должность.

Случай 3) отвечает уменьшению величины $\varkappa$ и увеличению χ, в случае 4) соответствующим образом изменяются функции p_1 и p_2, в случае 5) уменьшается число звеньев N в иерархии. Разумеется, могут быть реализованы и различные комбинации этих действий (в том числе и по отношению ко всей инстанции «заместители»). Если же начальник, наоборот, весьма удовлетворен каким-либо заместителем, то его действия будут, естественно, противоположными. Все эти изменения не затрагивают существа постулата и описываются с помощью изменения введенных выше функций p_1, p_2, $\varkappa$, χ, F. Соответствующая формализация должна, конечно же, проводиться тщательно и аккуратно.

3. Баланс власти в инстанции, условия на ее границах и переход к непрерывной модели. Теперь, когда введены основные модельные понятия, описаны взаимосвязи и предположения, можно сделать последний шаг в построении модели и ответить на вопрос: как найти распределение власти в иерархии, т. е. функцию $p(i,t)$? Для любого i-го звена властной структуры в любой момент времени t она подчиняется уравнению, при выводе которого используется своего рода локальный закон сохранения власти:

скорость изменения (уменьшения или увеличения) величины власти инстанции определяется потоками власти и реакцией системы.

Подсчитаем, какое количество власти получает (теряет) i-я инстанция за промежуток времени Δt между моментами t и $t + \Delta t$. Это количество формируется:

а) потоком власти, получаемым от $(i-1)$-й ступени по механизму (1), т. е. величиной

$$\Delta p_- = W(i-1,t)\,\Delta t;$$

б) потоком, отдаваемым в $(i+1)$-е звено по аналогичному механизму,

$$\Delta p_+ = -W(i,t)\,\Delta t;$$

в) суммой потоков, получаемых и отдаваемых по механизму (2) от удаленных инстанций $(j \neq i+1, i-1)$,

$$\Delta p_\Sigma = \sum_{j=0}^{N} V(i,j,t)\,\Delta t;$$

г) скоростью обмена властью между инстанцией и законодательством, которая определяется по предположению 2 реакцией гражданского общества,

$$\Delta p_F = F(i,t,p(i,t),p_1(i,t),p_2(i,t))\,\Delta t.$$

Складывая величины Δp_-, Δp_+, Δp_Σ, Δp_F, получаем суммарное изменение

$$\Delta p = p(i,t+\Delta t) - p(i,t) =$$

$$= \left[W(i-1,t) - W(i,t) + \sum_{j=0}^{N} V(i,j,t) + F(i,t,\ldots)\right]\Delta t. \qquad (3)$$

При выводе уравнения баланса (3) промежуток времени Δt предполагается, как обычно, достаточно малым, так чтобы величины W, V, F можно было считать постоянными. Используется также (и это существенно) предположение 1 о законопослушности, означающее, в частности, что в любом звене иерархии властное распоряжение всегда выполняется и поступившая с ним порция власти на деле реализуется (а не исчезает бесследно). Напомним, что введенный выше иерархический постулат уже нашел свое математическое воплощение в формулах (1), (2).

Разделив обе части (3) на Δt, получаем уравнение для скорости изменения власти i-й инстанции со временем

$$\frac{\Delta p(i,t)}{\Delta t} = -[W(i,t) - W(i-1,t)] + \sum_{j=0}^{N} V(i,j,t) + F(i,t,\ldots). \qquad (4)$$

Смысл (4) ясен из рассуждений, предшествовавших выводу баланса (3).

Уравнения (3), (4) записаны для произвольного номера $0 < i < N$. Для записи баланса в точке $i = 0$ необходимо определить величину $W(-1,t)$ потока власти, получаемой самой старшей инстанцией (с номером 0) от несуществующей вышестоящей инстанции. Очевидно, что

этот поток всегда равен нулю (высшей инстанции никто не может приказать). Аналогично, величина $W(N,t)$ — поток власти, передаваемый самой младшей ступенью иерархии в несуществующую нижестоящую ступень, также равен нулю (приказывать некому). Итак, поток власти на границах иерархии всегда равен нулю:

$$W(-1,t) = W(N,t) \equiv 0. \qquad (5)$$

Если рассматривается лишь часть иерархии, не включающая высшие и (или) низшие инстанции, то, очевидно, на границах должны задаваться (из тех или иных соображений) потоки власти как функции времени.

Для замыкания уравнения (4) необходимо, помимо условий (5), задать также распределение власти в некоторый начальный момент времени $t = t_0$, т. е. функцию

$$p(i,t_0) = p_0(i) \geqslant 0, \qquad 0 \leqslant i \leqslant N. \qquad (6)$$

Уравнение (4) с условиями (5), (6) и заданными функциями $\varkappa$, χ, F, p_1, p_2 представляет собой замкнутую дискретную модель распределения власти в системе «иерархическая структура—гражданское общество», математически реализующую общую схему, принятую в п. 1. Она позволяет для любых моментов времени $t > t_0$ и любых точек $0 \leqslant i \leqslant N$ найти решение (распределение власти в иерархии), т. е. функцию $p(i,t)$.

Осуществим теперь переход к непрерывной модели, рассматривая властную иерархию как «сплошную среду», т. е. считая число инстанций достаточно большим ($N \gg 1$), а все входящие в модель функции — достаточно гладкими. Дискретной координате i ставится в соответствие непрерывная координата x (номерам i соответствуют точки x_i, выбранные, например, по правилу $x_i = i$), а целочисленному отрезку $[0,N]$ — отрезок $[0,l]$ (величина l — аналог «длины» иерархической структуры, а координата x характеризует место инстанции в иерархии: чем больше x, тем младше инстанция). Разности вида $p(i+1,t) - p(i,t)$ в (1) и $p(i,t+\Delta t) - p(i,t)$, $W(i,t) - W(i-1,t)$ в (4) заменяются соответствующими первыми производными, сумма в правой части (4) — интегралом, а зависимость всех входящих в (4)–(6) функций от аргументов i, t заменяется на зависимость от x, t.

В результате для функции $p(x,t)$ (подробные выкладки не приводятся, так как они достаточно очевидны; ср. с выводом модели теплопередачи в § 2 гл. II) получается следующее уравнение:

$$\frac{\partial p}{\partial t} = \frac{\partial}{\partial x}\left[\varkappa\left(p,\frac{\partial p}{\partial x},x,t\right)\frac{\partial p}{\partial x}\right] +$$

$$+ \int_0^l \chi[p(x'),p(x),x',x]\cdot[p(x',t) - p(x,t)]\,dx' + F(p,p_1,p_2,x,t), \qquad (7)$$

$$0 < x < l, \quad t > t_0.$$

с граничными условиями второго рода

$$W(0,t) = -\varkappa \left.\frac{\partial p}{\partial x}\right|_{x=0} = 0, \quad W(l,t) = -\varkappa \left.\frac{\partial p}{\partial x}\right|_{x=l} = 0, \qquad t \geqslant t_0, \quad (8)$$

и начальными условиями

$$p(x,t_0) = p_0(x) \geqslant 0, \qquad 0 \leqslant x \leqslant l, \qquad (9)$$

причем $\varkappa > 0$, $\chi \geqslant 0$, а положительные функции p_1, p_2 монотонно убывают по x.

Уравнение (7) представляет собой «параболическое» интегро-дифференциальное уравнение (в том смысле, что в отсутствие интегрального члена оно аналогично уравнению теплопередачи; см. также упр. 1). В математическом отношении модель (7)–(9), «аппроксимирующая» исходную модель (4)–(6), замкнута и корректна, т. е. однозначно определяет решение — гладкую неотрицательную функцию $p(x,t)$ для всех $0 \leqslant x \leqslant l$ и $t \geqslant t_0$ (при некоторых несущественных для общности ограничениях на входные данные). Заметим, что при численном моделировании задачи (7)–(9) производится обратный переход к задаче (4)–(6) (см. упр. 2).

Напомним смысл величин, входящих в модель (7)–(9). Функция $p(x,t)$ (решение) описывает пространственно-временную динамику распределения власти в иерархической структуре, т. е. зависимость величины (уровня) реально осуществляемой инстанцией власти от ее местоположения (координата x) и времени t. Скорость изменения функции $p(x,t)$ (левая часть уравнения (7)) определяется следующими характеристиками:

1) разностью потоков власти, получаемой по механизму а) от ближайших соседей по иерархии или отдаваемой им (первый, дифференциальный член в правой части (7));

2) суммой потоков власти, получаемой инстанцией по механизму б) от отдаленных ступеней иерархии или отдаваемой им (второй, интегральный член в правой части (7));

3) реакцией гражданского общества, т. е. функцией

$$F(x,t,p(x,t),p_1(x,t),p_2(x,t))$$

(третий член в правой части (7));

4) минимальными и максимальными властными полномочиями иерархии — монотонно убывающими по x функциями $p_1(x,t) > 0$, $p_2(x,t) > 0$;

5) внутренними поведенческими свойствами иерархической структуры — функциями

$$\varkappa\left(p(x,t), \frac{\partial p}{\partial x}, x, t\right) > 0, \quad \chi(p(x',t), p(x,t), x', x, t) \geqslant 0;$$

6) начальным распределением власти в структуре — функцией $p_0(x) \geqslant 0$.

Таким образом, модель (7)–(9) — математическая реализация общей схемы системы «власть—общество», изложенной в п. 1 и отвечающей самоорганизующемуся объекту с различными прямыми и

обратными связями. Описываемое ею распределение власти в иерархии устанавливается не произвольно, а в результате взаимодействия всех элементов системы.

Очертим рамки применимости полученной модели. При ее построении существенно использовались постулат, предположение 1 о законопослушности системы, предположение 2 об определяющем влиянии реакции гражданского общества на скорость обмена властью между иерархией и законодательством, а также считалось, что характер текущего взаимодействия между двумя звеньями иерархии зависит от текущего состояния только этих инстанций.

Все эти предположения сформулированы в возможно наиболее общем виде, три из них представляются вполне естественными и обоснованными (что же касается предположения о законопослушности, то эта необходимая на начальном этапе исследования идеализация в дальнейшем может модифицироваться). Будем называть (7)–(9) *общей математической моделью динамики распределения власти* в государственных иерархических структурах, взаимодействующих с гражданским обществом.

Построенная выше модель позволяет анализировать ряд довольно общих проблем, связанных с функционированием и эволюцией системы «власть—общество». Перечислим некоторые из них:

1) получение необходимых и достаточных условий существования (или несуществования) стационарных (т. е. постоянных во времени) распределений власти и условий устойчивости таких стационаров;

2) анализ и прогноз кризисов власти различного характера: превышение власти; ненормальные, немонотонные распределения власти в иерархии, когда некоторые младшие инстанции осуществляют реальную власть, большую, чем некоторые старшие ступени; анархические или тоталитарные эволюции распределения власти и т. д.;

3) изучение результатов изменений, происходящих в реакции гражданского общества, в законодательстве, во взаимоотношениях внутри иерархии и т. д., на пространственно-временную динамику распределения власти.

Возможно также исследование различных комбинаций проблем 1)–3). Само собой подразумевается, что анализ перечисленных выше и других относящихся к существу дела задач должен сопровождаться тщательной их постановкой и аккуратной интерпретацией результатов в терминах используемой модели.

Модель (7)–(9) позволяет дать весьма общий ответ на ряд общих вопросов, возникающих при изучении проблем 1)–3). Например, используя теоремы сравнения решений параболических уравнений (см. § 2 гл. V), нетрудно установить, что при увеличении положительной реакции гражданского общества на действия всех инстанций власть в любой инстанции иерархии также увеличивается, а при увеличении отрицательной реакции — наоборот, а также что большие начальные распределения власти в иерархии приводят при прочих равных условиях к большему уровню власти во всех звеньях во все последующие моменты времени, и наоборот.

Однако первостепенный интерес представляют все же более конкретные вопросы, ответы на которые можно получить при определенных конкретизациях системы «власть—общество». Некоторые из них изучаются ниже.

4. Правовая система «власть—общество». Стационарные распределения и выход власти за рамки полномочий. Одна из наиболее естественно формализуемых систем «власть—общество» — правовая система. Поэтому простейшие конкретизации общей модели (7)–(9) отвечают именно этому случаю.

Система «власть—общество» называется *правовой*, если реакция общества на действия любой инстанции иерархии всегда направлена на удержание распределения власти в рамках предписанных ей полномочий. Подразумевается, что максимальные и минимальные полномочия соответствуют праву в общепринятом смысле этого понятия (подобный тип реакции отвечает правовому общественному сознанию).

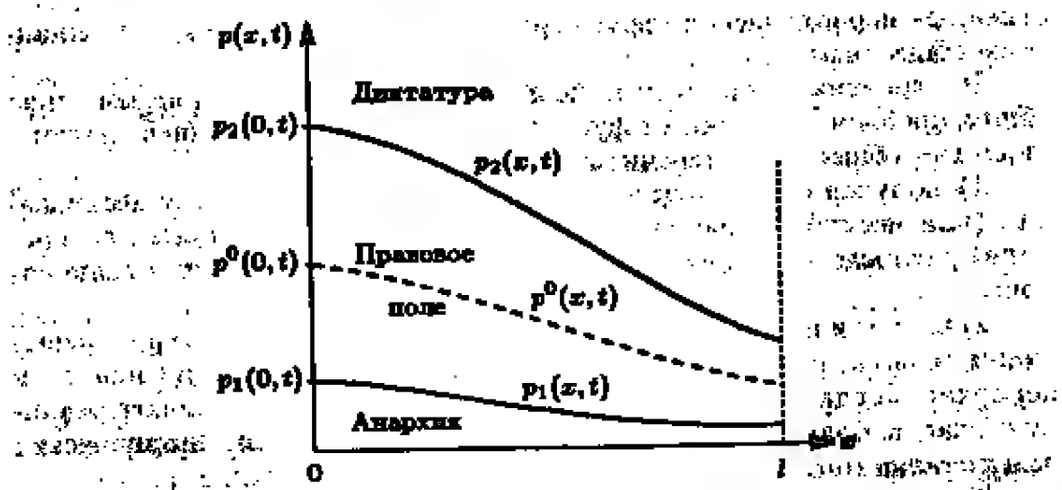


Рис. 63

В общей модели сформулированное выше определение реализуется заданием соответствующей функции $F(p, p_1, p_2, x, t)$. На рис. 63 показан качественный вид функций $p_2(x,t)$, $p_1(x,t)$ — максимальных и минимальных властных полномочий инстанций иерархии в некоторый момент времени t в зависимости от их места (координаты x). По своему смыслу $p_1(x,t)$, $p_2(x,t)$ — монотонно убывающие по x положительные функции, причем $p_2(x,t) > p_1(x,t)$, $0 \leqslant x \leqslant l$, $t \geqslant t_0$ (п. 1). Область, лежащая между кривыми p_1, p_2, — правовое поле, область выше значений p_2 условно называется диктатурой, область ниже значений p_1 — анархией. Искомое решение — распределение власти $p(x,t)$ — может априори в разные моменты времени либо полностью находиться в правовой области, либо частично или полностью выходить из нее.

Качественный вид зависимости функции $F(p, p_1, p_2, x, t)$ от величины власти $p(x,t)$ (в некоторый момент t для некоторого значения x) показан на рис. 64. Если величина власти превышает максимальные полномочия ($p > p_2$), то реакция общества отрицательна и растет с

ростом этого превышения; если же $p < p_1$, т. е. величина власти меньше минимальных полномочий, то, наоборот, реакция положительна и растет с увеличением отклонения. В промежутке $p_1 \leqslant p \leqslant p_2$ реакция общества с ростом p монотонно изменяется от положительных к отрицательным значениям, обращаясь в нуль в некоторой точке $p^0(x,t)$, $p_1 < p^0 < p_2$, $0 \leqslant x \leqslant l$, $t \geqslant t_0$.

Величину $p^0(x,t)$ можно назвать *идеальным* (средним) распределением власти в том смысле, что при реализации такого распределения реакция общества на действия иерархии всегда и всюду равнялась бы нулю (функция $p^0(x,t)$ считается монотонной по x; см. рис. 63).

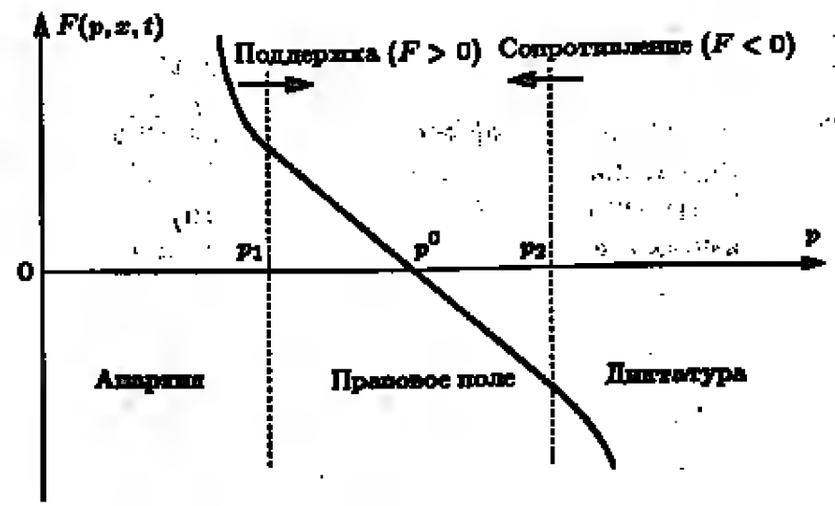


Рис. 64

Итак, в правовой системе общество отрицательно реагирует на положительное отклонение власти от некоторого среднего ее значения, усиливая величину реакции с увеличением этого отклонения (для отрицательных отклонений реализуется противоположная ситуация). По своей количественной амплитуде реакция зависит не только от разности $p - p^0$, но, вообще говоря, и от времени t; она может быть разной для разных звеньев иерархии (координата x) и может включать более тонкую зависимость от величины $p(x,t)$ (например, на рис. 64 интенсивность реакции для областей диктатуры $p > p_2$ и анархии $p < p_1$ заметно выше, чем для правовой области $p_1 < p < p_2$).

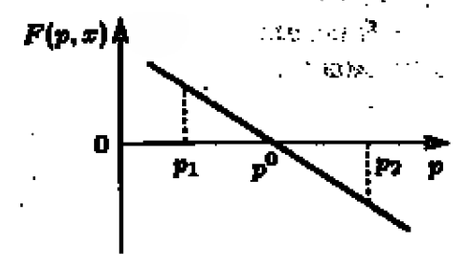


Рис. 65

Перейдем к изучению стационарных (не зависящих от времени) распределений власти в правовой системе. Начнем с анализа наиболее простой, но вместе с тем содержащей все существенные компоненты модели правовой системы. Для этого будем считать, что в общей модели (7)–(9):

1) функция $\varkappa$, соответствующая механизму передачи власти по команде, постоянна, т. е. $\varkappa = \varkappa_0 = \text{const} > 0$;

2) в иерархии нет механизма команд через голову, т. е. $\chi \equiv 0$;

3) реакция общества — линейная функция отклонения от идеального уровня власти и не зависит от t, т. е.

$$F(p,x) = k_1\,(p^0(x) - p),$$

где $k_1 > 0$ характеризует амплитуду реакции (рис. 65);

4) идеальное распределение власти p^0 и властные полномочия p_1, p_2 не меняются со временем и линейно убывают с ростом координаты x (рис. 66), т. е.

$$p^0 = H - kx, \quad H > 0, \quad p_1 = (1-\alpha)\,p^0, \quad p_2 = (1+\alpha)\,p^0.$$

Здесь $k > 0$ — степень убывания функции p^0 по x, $0 < \alpha < 1$ — относительная разница между p^0 и p_1, p_2 (причем $p^0 = (p_1 + p_2)/2$ — среднее арифметическое от минимальных и максимальных полномочий), $p^0(0) = H$, $p^0(l) = H - kl > 0$. Величина $(H - kl)/H = p^0(l)/p^0(0) = p_1(l)/p_1(0) = p_2(l)/p_2(0) < 1$ — отношение полномочий низших и высших степеней иерархии, или перепад полномочий.

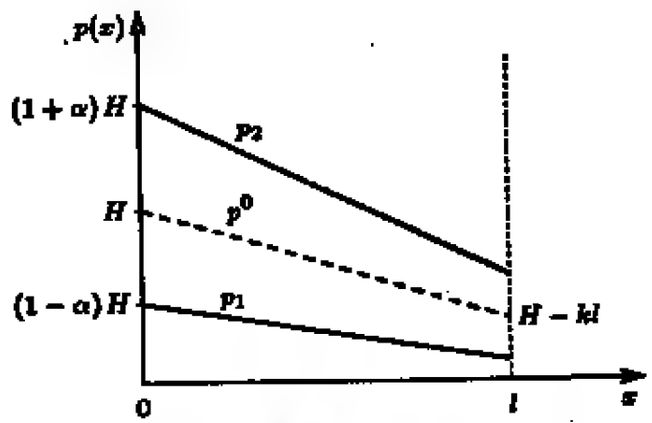


Рис. 66

В результате упрощений 1)–4) получаем *базовую модель* правовой системы:

$$\frac{\partial p}{\partial t} = \frac{\partial}{\partial x}\left(\varkappa_0 \frac{\partial p}{\partial x}\right) + k_1\,(p^0(x) - p), \quad 0 < x < l, \quad t > t_0,$$

$$\varkappa_0 \left.\frac{\partial p}{\partial x}\right|_{x=0} = \varkappa_0 \left.\frac{\partial p}{\partial x}\right|_{x=l} = 0, \quad t \geqslant t_0, \tag{10}$$

$$p(x, t_0) = p_0(x) \geqslant 0, \quad 0 \leqslant x \leqslant l,$$

где $k_1 > 0$, $p^0(x) = H - kx$, $H > 0$, $k > 0$.

Заметим, что модель (10) линейна, и для нее относительно нетрудно выписать общее решение.

Стационарные ($p(x,t) = p(x)$) решения (10) находятся из задачи

$$\varkappa_0 p'' = k_1\,[p - (H - kx)], \quad p'(0) = p'(l) = 0, \quad 0 < x < l, \quad p' = \frac{dp}{dx},$$

содержащей четыре параметра: $\varkappa_0$, k_1, k, l, и сводящейся масштабированием (см. § 1 гл. V) $\bar{p} = p/H$, $\bar{x} = x/l$ к задаче с двумя параметрами

$$ap'' = p - (1 - bx), \quad 0 < x < l, \quad p'(0) = p'(1) = 0, \tag{11}$$

где $a = \varkappa_0/(k_1 \cdot l^2)$, $b = kl/H < 1$ (черточки над величинами p и x в (11) опущены).

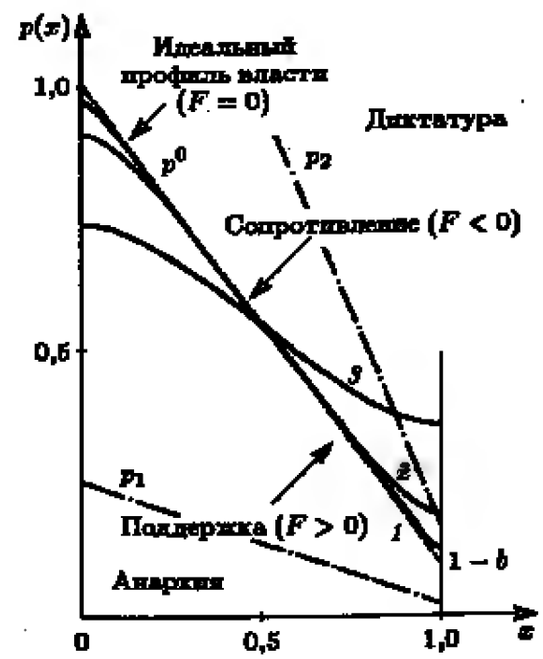


Рис. 67

Задача (11) имеет хорошо известное единственное (неотрицательное) решение, даваемое монотонно убывающей по x функцией

$$p(x) = \frac{a^{1/2} b}{e^{a^{-1/2}} - e^{-a^{-1/2}}}\left[\left(1 - e^{-a^{-1/2}}\right) e^{xa^{-1/2}} + \left(1 - e^{a^{-1/2}}\right) e^{-xa^{-1/2}}\right] + (1 - bx). \tag{12}$$

Качественное поведение решения (12) определяет параметр a (параметр $b \approx 1$, так что $1 - b \ll 1$, поскольку рассматриваются достаточно протяженные иерархические структуры с заметным перепадом полномочий). Вид решения для значений $b = 0{,}9$ и трех значений $a = 10^{-3}$, 10^{-2}, 10^{-1} приведен на рис. 67. С ростом a решение (кривые *1–3*) становится более пологим, и при некотором $a > a_{кр}$ решение частично лежит в области $p > p_2$ (кривая *3*). Анализируя (12) и рис. 67, можно сделать следующие предварительные выводы:

1) стационарное распределение власти в правовой системе существует;

2) существует область параметров системы, в которой распределение власти выходит за рамки полномочий (в данном случае для части младших ступеней власть превышена. Нетрудно выписать необходимое и достаточное условие нахождения власти в правовой области:

$$a^{1/2}b\left(e^{a^{-1/2}}-1\right)\Big/\left(e^{a^{-1/2}}+1\right)\leqslant(1-b)(1+\alpha). \qquad (13)$$

С ростом величины a критерий (13) нарушается все сильнее, причем если $a\to\infty$, то $p(x)\to 1-b/2$. Максимальное превышение (достигаемое в точке $x=1$) равно $p(1)-p_2(1)=1-b/2-(1+\alpha)(1-b)$, что при $b\approx 1$ дает $p(1)-p_2(1)\approx 0{,}5$, т. е. превышение составляет 50% от максимальных полномочий;

3) идеальное распределение власти (штриховая прямая на рис. 67) не реализуется, решение всегда имеет «добавку» к величине $p^0(x)=1-bx$ — второму члену в правой части (12).

Ввиду важности выводов 1), 2) необходимо выяснить степень их общности, т. е. понять, не являются ли они свойствами только упрощенной модели (10).

5. Роль основных характеристик системы в феномене превышения (принижения) власти. Введем некоторые обобщения в базовую модель (10).

1) *Механизм команд через голову.* В правой части уравнения (10) появляется дополнительное слагаемое вида

$$\int_0^1 \chi[p(x')-p(x)]\,dx'$$

(здесь взято $\chi=\chi_0=\mathrm{const}>0$). Задача для стационарного решения (аналог задачи (11)) после соответствующего масштабирования имеет вид

$$ap''=[p-(1-bx)]-\sigma\int_0^1[p(x')-p(x)]\,dx', \qquad (14)$$

$$p'(0)=p'(1)=0, \qquad 0<x<1,$$

где $\sigma=\chi_0 l/k_1>0$. Несложной заменой (14) сводится к задаче (11) с ненулевыми краевыми условиями (упр. 3), имеющей решение вида (12). Условие (13) переходит в условие

$$\frac{b}{1+\sigma}\left[z^{-1/2}\frac{e^{z^{1/2}}-1}{e^{z^{1/2}}+1}+\frac{\sigma}{2}\right]\leqslant(1-b)(1+\alpha), \qquad z=\frac{1+\sigma}{a},$$

которое так же, как и (13), всегда нарушается при достаточно больших значениях a и (или) σ (максимальное превышение то же, что и в случае $\chi_0=0$, так как $p(x)\to 1-b/2$ при больших a, σ).

2) *Уточненная реакция гражданского общества.* В обществе существует часть, которая в принципе не согласна с любыми действиями иерархии (в базовой модели как бы предполагалось, что такой компоненты нет). Подобная реакция может быть описана добавлением в правую часть уравнения (10) слагаемого вида k_2p, $k_2<0$ (т. е. эта часть реакции всегда отрицательна и тем больше, чем больше величина власти). Для функции $p(x)$ в этом случае получается задача, аналогичная (11). Однако при достаточно больших значениях k_2 и a решение, очевидно, уходя из области $p>p_2$, попадает в область $p<p_1$, т. е. распределение власти выходит за рамки минимальных полномочий (в данном случае это происходит с высшими инстанциями).

3) *Различные механизмы передачи власти по команде.*

а) В первом члене правой части уравнения (10) рассматривается не функция $\varkappa=\varkappa_0=\mathrm{const}$, а функция $\varkappa=\varkappa_0|\partial p/\partial x|^\alpha$, $\alpha>-1$. Это значит, что степень ответственности иерархии зависит от градиента величины $p(x,t)$ (неравенство $\alpha>-1$ обеспечивает параболичность соответствующего уравнения). Для получающейся стационарной задачи существует единственное монотонное решение. При достаточно больших значениях параметра, аналогичного параметрам a и σ из (11), (14), распределение власти выходит за рамки полномочий (поскольку с его ростом решение стремится к некоторой константе).

б) В первом члене правой части уравнения (10) вместо функции $\varkappa=\varkappa_0=\mathrm{const}$ берется функция $\varkappa=\varkappa_0p^\beta$, $\beta>-1$, т. е. ответственность звеньев иерархии зависит от величины реализуемой ими власти. По сравнению с предыдущими случаями результаты принципиально не меняются.

В случаях а), б) эти выводы получаются из анализа полей интегральных кривых соответствующих нелинейных уравнений первого порядка (см. упр. 4).

6. Интерпретация результатов и выводы. Итак, из 1)–3) п. 5 следует, что выводы 1), 2) п. 4 о существовании стационарных распределений власти в правовой системе и наличии феномена превышения (принижения) власти (даже в такой системе) носят достаточно общий характер. Если вывод 1) п. 4 безусловно положительный, то вывод 2) п. 4 необходимо обсудить более подробно, поскольку он означает наличие явного кризиса власти.

Во всех случаях, рассмотренных в пп. 4, 5, выход власти за рамки полномочий происходит при достаточно больших значениях некоторых параметров a и (или) σ (или их аналогов). Чтобы понять суть проблемы, обратимся к базовой модели п. 4. В задаче (11) параметр a определяется формулой

$$a=\frac{\varkappa_0}{k_1l^2}, \qquad (15)$$

являясь *системным параметром* (содержащим несколько характеристик объекта), а также *параметром подобия* (для систем с различ-

ными значениями $\varkappa_0$, k_1, l, но с одинаковыми значениями a профиль власти имеет один и тот же качественный вид).

Из (15) видно, что a увеличивается с ростом параметра $\varkappa_0$, характеризующего интенсивность механизма передачи власти по команде, и с уменьшением параметров k_1, l — реакции общества и длины иерархической структуры.

Пусть k_1, l фиксированы, а $\varkappa_0$ растет (вместе с ним растет и величина a). С ростом $\varkappa_0$ уменьшается ответственность всех звеньев иерархии в том смысле, что старшие инстанции все более легко «транспортируют» потоки власти младшим соседям (см. п. 2). Предельная безответственность ($\varkappa_0 = \infty$) означает, что любая инстанция становится своего рода телетайпом, передающим полученные «сверху» властные распоряжения (и поступающие с ними порции власти) «вниз» без всякого их изменения. В результате у нижних ступеней иерархии как бы накапливаются излишки власти (кривая *3* на рис. 67), которые не могут быть уменьшены за счет реакции общества (она фиксирована). Поэтому низшие инстанции вынуждены выйти за рамки своих максимальных полномочий. Аналогичная картина наблюдается при изменении параметра $\sigma = \chi_0 l/k_1$ (команды через голову).

Бюрократический смысл условия (13) и его аналогов заключается в том, что, если характеристики иерархии и гражданского общества известным образом согласованы, то распределение власти остается в правовой области (и наоборот).

Рассуждения, сопровождающие анализ соотношения (15), служат выражением общего противоречия, присущего иерархической структуре. Оно состоит в следующем. С одной стороны, в иерархии обязательно должен быть перепад полномочий (причем весьма значительный), соответствующих ее высшим и низшим звеньям, иначе иерархия не имеет смысла. С другой стороны, механизмы передачи власти внутри иерархии всегда работают на сглаживание разницы во власти между инстанциями. Если данные механизмы становятся все сильнее, то остающаяся без изменений реакция гражданского общества даже в правовой системе не в состоянии конкурировать с этими процессами и скомпенсировать сглаживание. Распределение власти становится все более пологим и с необходимостью частично выходит за рамки предписываемых ей полномочий. Заметим, что данное свойство имеет место не только для общей правовой системы «власть—общество».

Таким образом, справедливы два следующих вывода:

1) для правовой системы всегда существует единственное стационарное распределение власти в иерархической структуре, причем величина власти монотонно падает при переходе от старших к младшим инстанциям;

2) даже в правовой системе всегда существует область параметров, при реализации которых распределение власти выходит за границы правового поля.

Столь же общим является и вывод 3) из п. 4 о невозможности реализации идеального распределения власти, т. е. такого распределения, при котором реакция общества всегда и всюду была бы нулевой. Этот вывод устанавливается элементарно, и его смысл достаточно ясен: при

нулевой реакции иерархия перестала бы взаимодействовать со своим партнером — гражданским обществом. В то же время согласно модели иерархическая цепочка подвержена противоположно направленным влияниям (поддержка и сопротивление общества, «прибыль» и «убыль» власти в инстанциях). В условиях правовой системы их динамическая конкуренция формирует в итоге стационарный профиль власти подобно тому, как статическая конкуренция сил тяжести и упругости определяет форму троса, висящего между опорами. В отсутствие реакции единственно возможным решением в этом случае было бы постоянное по пространству и во времени распределение власти (упр. 4). Из этого вывода, в частности, следует, что при уменьшении разницы между максимальными и минимальными полномочиями (в предельном случае это означает равенство величин $p_1 \equiv p_2 \equiv p^0$) распределение власти всегда выходит из правовой области. Другими словами, для действий иерархии не следует отводить слишком узкое правовое поле, иначе она будет «вынужденно» выходить за законные рамки своей власти.

При построении рассмотренной модели системы «власть—общество» и ее обобщений широко использовались аналогии, почерпнутые из изучения естественнонаучных объектов: переход от дискретной к непрерывной модели (иерархия трактуется как сплошная среда); аналог закона Фурье при описании механизмов перераспределения власти в иерархии; «закон сохранения власти» при выводе основного уравнения; понятие «поток власти» и др. Эти аналогии допускают разумную интепретацию, их применение позволяет глубже понять принципиальные свойства столь трудноформализуемого объекта.

УПРАЖНЕНИЯ

1. Уравнение вида $u_t = L(u_{xx}, u_x, u, x, t)$ называется *параболическим*, если $\partial L/\partial u_{xx} > 0$. Получите условия для функции $\varkappa(p, p_x, x, t)$ в уравнении (7), записанном без интегрального члена, при которых оно является параболическим.

2. Положив $\chi = 0$, осуществите, используя разложение функций в ряд Тейлора и удерживая главные члены, переход от модели (4)–(6) к (7)–(9). Заменяя производные конечными разностями, выполните обратный переход от модели (7)–(9) к (4)–(6).

3. Проведя замену $v(x) = p(x)(1+\sigma) - (1-bx) - \sigma B$, где $B = \int_0^1 p(x')\,dx'$, получите из (14) краевую задачу для функции $v(x)$.

4. Представляя стационарное уравнение (10) в виде системы двух уравнений, найдите для случаев $\varkappa = \varkappa_0$, $\varkappa = \varkappa_0 |\partial p/\partial x|^\alpha$, $\alpha > -1$ и $\varkappa = \varkappa_0 p^\beta$, $\beta > -1$ замены, сводящие его к уравнению первого порядка.

5. Проверьте, что при $F \equiv 0$, $\chi \equiv 0$ решение уравнения (7) с условиями (8) в стационарном случае имеет вид $p(x) = \mathrm{const}$.

Библиография к главе IV: [1, 5, 9, 10, 37, 38, 40, 48–50, 56, 58, 85, 86].

ГЛАВА V

ИССЛЕДОВАНИЕ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ

§ 1. Применение методов подобия

Дадим характеристику способов упрощения математических моделей, основанных на свойствах их симметрии. Приведем описание самоподобных (автомодельных) явлений. Исследуем некоторые автомодельные процессы для нелинейных параболических и гиперболических уравнений.

1. Анализ размерностей и групповой анализ моделей. Одно из фундаментальных свойств природных, технологических, многих экономических и социальных объектов — симметрия (подобие, повторяемость, воспроизводимость) — находит свое отражение в их математических моделях. Наличие какого-либо вида симметрии у изучаемого явления означает большую простоту объекта в сравнении с его менее симметричным аналогом. На этом основываются широко применяемые методы упрощения математических моделей и, следовательно, методы упрощения их анализа. Они состоят в понижении порядка системы уравнений, образующих модель, в уменьшении числа переменных, от которых зависят искомые величины, или числа постоянных параметров, определяющих процесс, и т. д. (так, симметрия функции трех переменных относительно этих переменных позволила найти максимально возможное значение скорости трехступенчатой ракеты; см. п. 4 § 1 гл. I).

Типичный подход к использованию свойств симметрии — анализ размерности величин, входящих в модель. Часть характеристик объектов измеряется в каких-либо единицах, имеющих непосредственный (механический, физический, экономический и т. д.) смысл. Например, масса в граммах, температура в градусах Кельвина, валовый национальный продукт в рублях. Такие величины называются *размерными*, их численное значение зависит от выбора единиц измерения. Среди них выделяются величины с независимой (основной) размерностью, или *размерно независимые* величины. Например, если для описания механических явлений используется система единиц СГС (сантиметр, грамм, секунда), то размерности длины x, массы m и времени t независимы и не выражаются одна через другую. В отличие от них размерность кинетической энергии $E = mv^2/2$ определяется через размерности основных величин по формуле $[E] = [m][x]^2[t]^{-2} = \text{г}\cdot\text{см}^2\cdot\text{с}^{-2}$,

называемой *формулой размерности* (здесь $v = dx/dt$, через символ $[f]$ обозначается размерность величины f). Такие величины называются *размерно зависимыми*. Напомним, что явления и процессы могут описываться также и безразмерными величинами, скажем, отношением длины водоносного пласта к его ширине, показателем степени в формуле, дающей зависимость коэффициента теплопроводности от температуры, годовым банковским процентом и т. п.

Системы единиц измерения можно выбирать по-разному, причем связи между величинами, характеризующими объект (получаемые из законов природы или иных соображений), не должны изменяться при изменении единиц измерения. Например, второй закон Ньютона $F = ma$ (F — сила, a — ускорение) в системе СИ записывается точно так же, как и в системе СГС. Инвариантность явлений и процессов по отношению к изменению единиц измерения находит свое воплощение в так называемой П-теореме.

Пусть имеется функциональная связь

$$a = F(a_1, a_2, \ldots, a_k, a_{k+1}, \ldots, a_n) \tag{1}$$

между $n+1$ *размерными величинами* a, a_1, ..., a_n, *где величины* a_1, ..., a_k *имеют независимую размерность, и пусть эта связь не зависит от выбора системы единиц измерения (величина* a *искомая, а остальные задаваемые).*

Тогда связь (1) *может быть записана как*

$$\Pi = F(\underbrace{1, \ldots, 1}_{k}, \underbrace{\Pi_1, \ldots, \Pi_{n-k}}_{n-k}), \tag{2}$$

т. е. в виде соотношения между $n+1-k$ *величинами* Π, Π_1, ..., Π_{n-k}, *представляющими собой безразмерные комбинации из* $n+1$ *размерных величин* a, a_1, ..., a_n.

При этом величины Π, Π_1, ..., Π_{n-k} связаны с a, a_1, ..., a_n простыми соотношениями

$$\begin{aligned} a &= \Pi a_1^{m_1} a_2^{m_2} \ldots a_k^{m_k}, \\ a_{k+1} &= \Pi_1 a_1^{l_1} a_2^{l_2} \ldots a_k^{l_k}, \\ &\ldots\ldots\ldots\ldots \\ a_n &= \Pi_{n-k} a_1^{p_1} a_2^{p_2} \ldots a_k^{p_k}. \end{aligned} \tag{3}$$

Здесь показатели степеней m_1, ..., m_k; l_1, ..., l_k; ...; p_1, ..., p_k те же, что и в соответствующих формулах размерностей для размерно зависимых величин a, a_{k+1}, ..., a_n, например в формуле $[a] = [a_1]^{m_1}[a_2]^{m_2}\ldots[a_k]^{m_k}$.

Доказательство П-теоремы основано на инвариантности связи (1) относительно единиц измерения. Прежде всего проведем *обезразмеривание* соотношения (1), исходя их того, что любая размерно независимая величина a_i, $i = 1, \ldots, k$, может быть представлена в виде $a_i = \bar{a}_i a_i$. Здесь безразмерный коэффициент $\bar{a}_i$ — числовое значение величины a_i

в применяемой системе единиц, а сомножитель α_i имеет размерность a_i и характеризует масштаб измерения (десятки или сотни фунтов, сотни или тысячи градусов, миллионы или миллиарды рублей и т. д.). Числовые значения безразмерных множителей для размерно зависимых величин a, $a_{k+1}, \ldots, a_n$ вычисляются с использованием *масштабных множителей* α_i, $i = 1, \ldots, k$, по правилу

$$\bar{a} = \frac{a}{\alpha_1^{m_1}\alpha_2^{m_2}\ldots\alpha_k^{m_k}}, \quad \bar{a}_{k+1} = \frac{a_{k+1}}{\alpha_1^{l_1}\alpha_2^{l_2}\ldots\alpha_k^{l_k}}, \quad \bar{a}_n = \frac{a_n}{\alpha_1^{p_1}\alpha_2^{p_2}\ldots\alpha_k^{p_k}},$$

непосредственно следующему из формул размерностей для каждой из них. Соотношение (1) можно трактовать также и как связь между числовыми значениями величин a, $a_1, \ldots, a_n$ (т. е. связь между безразмерными величинами $\bar{a}, \bar{a}_1, \ldots, \bar{a}_n$), не зависящую, по предположению, от единиц измерения. Таким образом, для любых наборов масштабных множителей α_i справедливо

$$\bar{a} = F(\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_k, \bar{a}_{k+1}, \ldots, \bar{a}_n),$$

или

$$\frac{a}{\alpha_1^{m_1}\alpha_2^{m_2}\ldots\alpha_k^{m_k}} = F\left(\frac{a_1}{\alpha_1}, \frac{a_2}{\alpha_2}, \ldots, \frac{a_k}{\alpha_k}, \frac{a_{k+1}}{\alpha_1^{l_1}\alpha_2^{l_2}\ldots\alpha_k^{l_k}}, \ldots, \frac{a_n}{\alpha_1^{p_1}\alpha_2^{p_2}\ldots\alpha_k^{p_k}}\right).$$

Положим теперь $\alpha_1 = a_1$, $\alpha_2 = a_2$, $\ldots$, $\alpha_k = a_k$. Другими словами, выберем масштабные множители так, чтобы в полученной системе единиц измерения величины $\bar{a}_1, \ldots, \bar{a}_k$ тождественно равнялись единице. Тогда из последнего соотношения немедленно вытекают формулы (2) и (3).

Применение П-теоремы снижает число величин, фигурирующих в описании объекта, и дает явный способ представления искомой величины a (и величин $a_{k+1}, \ldots, a_n$) через Π, $\Pi_1, \ldots, \Pi_{n-k}$ и $a_1, \ldots, a_k$. Он «безразличен» к конкретному виду функциональной зависимости (1), требуется лишь достаточная гладкость функции F.

В частности, если $n = k$, то, как сразу следует из (2), $\Pi = \text{const}$, и

$$a = \text{const} \cdot a_1^{m_1} a_2^{m_2} \ldots a_k^{m_k},$$

т. е. для решения получается простое выражение через задаваемые параметры (чтобы знать точное значение a, остается определить константу). Пусть, например, известно, что период колебания маятника T (см. п. 3 § 3 гл. I) не зависит от его начального отклонения и скорости, а определяется лишь его длиной l, массой m и ускорением свободного падения g. Функциональная связь $T = T(l, m, g)$ содержит четыре размерных величины, три из которых имеют независимые размерности. Выберем в качестве таковых T, l и m, тогда для размерности g будем иметь $[g] = [l][T]^{-2}$, или $[T] = [l]^{1/2}[g]^{-1/2}$, откуда

$$T = \text{const} \cdot \sqrt{l/g}.$$

С точностью до безразмерного множителя данная формула совпадает с полученной из решения уравнения колебаний маятника (попутно выясняется, что их период не зависит от m).

Заметим, что безразмерные параметры, характеризующие объект, при вариации единиц измерения не изменяются и поэтому в П-теореме не фигурируют.

Процедура обезразмеривания (*масштабирования*) всегда полезна при изучении математических моделей, поскольку может дать важную предварительную информацию об объекте. Например, в результате масштабирования задачи п. 4 § 4 гл. IV выяснилось, что ее решение определяется фактически не четырьмя, а лишь одним параметром.

Получаемые с помощью П-теоремы безразмерные величины $\Pi_1, \ldots, \Pi_{n-k}$ можно назвать *параметрами* (*критериями*) *подобия* в том смысле, что разные по своим масштабам, но одинаковые по сущности явления и процессы ведут себя качественно одинаково при заданном наборе параметров $\Pi_1, \ldots, \Pi_{n-k}$ (и одинаково изменяются при их изменении).

Инвариантность моделей по отношению к системе единиц измерения — частный случай более общих свойств их симметрии. Наиболее хорошо разработанный и широко применяемый подход, использующий подобие моделей, основан на так называемом *инвариантно-групповом методе* исследования дифференциальных уравнений. Действительно, большинство дифференциальных уравнений, представляющих собой составную часть математических моделей многих явлений, остаются неизменными (*инвариантными*) при некоторых преобразованиях входящих в них независимых переменных и искомых функций.

Например, уравнение теплопередачи (5) из § 2 гл. II

$$c\frac{\partial T}{\partial t} = \operatorname{div}(\varkappa \cdot \operatorname{grad} T) \tag{4}$$

инвариантно к «сдвигу» времени $t' = t + t_0$ и, если функции c и $\varkappa$ не зависят от $\vec{r}$, к «сдвигу» координат $\vec{r}' = \vec{r} + \vec{r}_0$. В частном случае $c = c_0$, $\varkappa = \varkappa_0 T^\sigma$, т. е. при степенной зависимости коэффициента теплопроводности от температуры, (4) не меняет своего вида при преобразованиях «растяжения—сжатия»: $t' = \alpha t$, $\vec{r}' = \beta\vec{r}$, $T' = \gamma T$ (числа α, β, γ должны подчиняться некоторой связи; см. упр. 1).

Аналогичными свойствами обладают, как легко видеть, одномерные уравнения газовой динамики для идеального политропного газа

$$\frac{\partial}{\partial t}\left(\frac{1}{\rho}\right) = \frac{\partial v}{\partial m}, \quad \frac{\partial v}{\partial t} = -\frac{\partial p}{\partial m}, \quad \frac{\partial}{\partial t}(p\rho^{-\gamma}) = 0, \tag{5}$$

записанные в массовых координатах (см. пп. 5, 7 § 4 гл. II). Нетрудно убедиться, что в том или ином виде инвариантность присуща большинству построенных в гл. I–IV моделей.

Рассмотренные выше преобразования переменных и различные их комбинации и обобщения относятся к классу так называемых точечных преобразований, или преобразований Ли (производные от величин

не преобразуются). Пусть они к тому же удовлетворяют некоторым дополнительным условиям (представляют собой группу). Тогда дифференциальное уравнение, допускающее группу Ли, может быть упрощено: либо понижается его порядок, либо уменьшается число независимых переменных, определяющих искомые функции. Одна из целей группового анализа — определение всех групп преобразований, допускаемых данным уравнением или системой уравнений, и соответствующих этим группам так называемых *инвариантных решений* уравнений. С точки зрения исследования математических моделей важны, прежде всего, не сложная и громоздкая процедура группового анализа, а его конечные итоги применительно к конкретному уравнению (для большинства базовых математических моделей эти результаты получены).

Приведем в качестве важного частного примера итог группового анализа уравнения (4) при $c = 1$, $\varkappa = \varkappa(T)$ в одномерном случае (табл. 3). Номера в левом столбце таблицы отвечают различным групповым преобразованиям, а в ее клетках приведены соответствующие им инвариантные решения для различных видов (спецификаций) функ-

Таблица 3

Групповая классификация уравнения

$$\frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(\varkappa(T)\frac{\partial T}{\partial x}\right)$$

№	$\varkappa(T)$ произвольное	$\varkappa(T) = e^T$	$\varkappa(T) = T^\sigma$ $(\sigma \neq -4/3)$
1	$T = f(\xi)$, $\xi = x$	*	*
2	$T = f(\xi)$, $\xi = t$	*	*
3	$T = f(\xi)$, $\xi = x - t$	*	*
4	$T = f(\xi)$, $\xi = x^2/t$	*	*
5	—	$T = \frac{1}{\alpha}\ln x + f(\xi)$, $\xi = tx^{\frac{1-2\alpha}{\alpha}}$; $T = \ln t + f(\xi)$, $\xi = x$ $(\alpha = 0)$	$T = t^{1/(2\alpha)} f(\xi)$, $\xi = xt^\beta$, $\beta = -(\alpha + \sigma/2)/(2\alpha)$; $T = t^{-1/\sigma} f(\xi)$, $\xi = x$ $(\alpha = -\sigma/2)$
6	—	$T = x + f(\xi)$, $\xi = te^x$	$T = e^t f(\xi)$, $\xi = xe^{-\sigma t/2}$
7	—	$T = 2t + f(\xi)$, $\xi = xe^{-t}$	$T = e^x f(\xi)$, $\xi = te^{\sigma x}$

ции $\varkappa(T)$. Символ * означает повтор выражения, стоящего в строке слева, прочерк — отсутствие инвариантных решений. Через ξ обозначен инвариант — комбинация, не изменяющаяся при данном преобразовании.

При подстановке инвариантного решения $T(x,t)$ в одномерное уравнение (4) получается обыкновенное дифференциальное уравнение относительно функции $f(\xi)$ (также инвариант преобразования). Исследовать его гораздо проще, чем исходное уравнение в частных производных. Далее не представляет труда выписать функцию $T(x,t)$ через $f(\xi)$ и изучить ее свойства, если известны свойства $f(\xi)$.

Групповые преобразования с номерами 1–4 основные, поскольку допускаются при произвольных спецификациях функции $\varkappa(T)$. Случай 1 — стационарное решение (инвариантность (4) по отношению к сдвигу по t), 2 — постоянное по пространству решение (инвариантность к сдвигу по x), 3 — решение типа бегущей волны (инвариантность к одновременному сдвигу по t и x), 4 — решение с постоянным значением функции $T(x,t)$ в точке $x = 0$ (инвариантность к преобразованиям вида $t' = \alpha^2 t$, $x' = \alpha x$, $T'(x',t') = T(x,t)$).

Расширения основной группы (называемые *подгруппами*) происходят при конкретизациях $\varkappa(T) = e^T$, $\varkappa(T) = T^\sigma$ функции $\varkappa(T)$. Так, для решений пятой подгруппы величина ξ инвариантна к преобразованию растяжения-сжатия, которое в этом случае допускается исходным уравнением (поясним: решение 4 отвечает фактически частному случаю преобразований растяжения-сжатия для произвольной функции $\varkappa(T)$, а для $\varkappa(T) = T^\sigma$ им отвечают решения в разделяющихся переменных). Решения для подгрупп 6 и 7 порождаются комбинацией из преобразований сдвига и растяжения-сжатия.

Заметим: вид почти всех инвариантных решений из табл. 3 был получен еще до группового анализа из других соображений (из теории размерностей, предельными переходами, угадыванием решения и прямой подстановкой его в уравнение (4)). Однако это не умаляет достоинства данного метода, по крайней мере в силу двух причин. Он дает строгий и исчерпывающий ответ об основной группе преобразований и о тех спецификациях уравнений (моделей), при которых эта группа расширяется (заранее известно, что другие спецификации $\varkappa(T)$ в уравнении (4), скажем, вида $\varkappa(T) = \ln(1+T)$, не допускают дополнительных к основным решений). Кроме того, свойства симметрии уравнений могут быть весьма разнообразными и неожиданными. Например, для не описанного в табл. 3 случая $\varkappa = T^\sigma$, $\sigma = -4/3$ уравнение (4) имеет большое число нетривиальных преобразований, определяемых лишь с помощью группового анализа, представляющего собой регулярный способ упрощения математических моделей.

Подчеркнем также, что этот метод изучает свойства дифференциального уравнения как такового, т. е. свойства лишь части модели исходного объекта (без других входных данных, например без краевых условий). Поэтому пригодность получаемых на его основе тех или иных инвариантных решений для описания конкретного явления должна исследоваться дополнительно. В отличие от группового анализа, теория размерностей (ее можно рассматривать как примене-

ние частного случая групповой симметрии, сводящейся к растяжению-сжатию единиц измерения) и вытекающая из нее П-теорема оперируют с моделью во всей ее полноте.

2. Автомодельные (самоподобные) процессы. Среди инвариантных решений дифференциальных уравнений выделяется важный класс *самоподобных*, или *автомодельных* решений (смысл этого названия станет понятным чуть ниже). К ним принято относить широко используемые решения типа бегущей волны (случай 4 из табл. 3 для $\varkappa = T^a$; см. также упр. 4 § 2 гл. II), степенные автомодельные решения (случай 5; см. также упр. 3 § 2 гл. II) и экспоненциальные автомодельные решения (случаи 6, 7).

Убедимся в полезности построения и анализа автомодельных решений для исследования математических моделей. В качестве первого примера рассмотрим решения типа бегущей волны для системы уравнений (5). Они ищутся в виде

$$\rho(m,t) = \rho(\xi) = \rho(m - Dt),$$
$$v(m,t) = v(\xi) = v(m - Dt),$$
$$p(m,t) = p(\xi) = p(m - Dt),$$

где $D > 0$ — некоторая постоянная. Подставляя эти выражения в (5) с заменой для единообразия третьего уравнения дивергентным уравнением энергии (21) § 4 гл. II, получим вместо уравнений в частных производных систему трех обыкновенных дифференциальных уравнений

$$D\frac{d}{d\xi}\frac{1}{\rho} + \frac{dv}{d\xi} = 0, \quad -D\frac{dv}{d\xi} + \frac{dp}{d\xi} = 0, \quad -D\frac{d}{d\xi}\left(\varepsilon + \frac{v^2}{2}\right) + \frac{d}{d\xi}(pv) = 0,$$

где $\varepsilon(\xi) = \varepsilon(m - Dt)$ — внутренняя энергия газа ($\varepsilon = \varepsilon(p,\rho)$). Интегрируя их в произвольных пределах ξ_0, ξ_1, убеждаемся в том, что бегущая волна отвечает течениям, в которых сохраняются три интеграла:

$$D\frac{1}{\rho} + v = C_\rho, \quad -Dv + p = C_p, \quad D\varepsilon + D\frac{v^2}{2} - pv = C_\varepsilon. \qquad (6)$$

В случае непрерывных течений интегралы (6) дают единственное решение — постоянные при всех ξ значения ρ, v, p, ε.

Нетривиальный результат получается, если учесть, что уравнения газовой динамики благодаря существованию в них «градиентной катастрофы» могут допускать *разрывные решения* (см. п. 7 § 4 гл. II). Пусть разрыв функций $\rho(\xi)$, $v(\xi)$, $p(\xi)$ расположен в точке $\xi = 0$. В областях $\xi > 0$, $\xi < 0$, т. е. справа и слева от него, течение, очевидно, постоянно и характеризуется наборами величин ρ_0, v_0, p_0 и ρ_1, v_1, p_1 (см. рис. 68, на котором решение изображено как функция координаты m в моменты времени t_1, t_2, t_3; $t_1 < t_2 < t_3$). Скачок газодинамических параметров не произволен. Покажем это, анализируя смысл написанных выше интегралов (справедливых также и для разрывных течений).

Описываемый автомодельным решением разрыв движется по массе газа с постоянной скоростью D. Неизменность массовой скорости

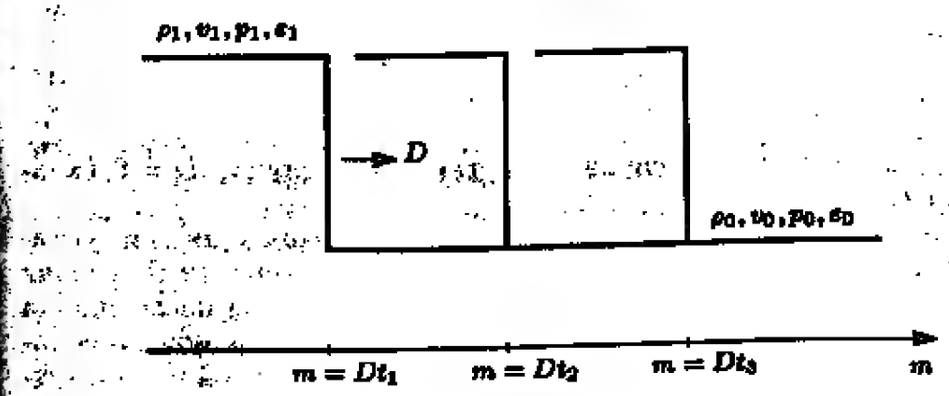


Рис. 68

обеспечивает равенство потока вещества $I_\rho = D$, «втекающего» с одной стороны разрыва и «вытекающего» с другой его стороны. Нарушение этого естественного физического требования означало бы появление или исчезновение при переходе через него какого-то количества вещества. Поток «втекающей» массы равен, по определению, $I_{\rho_0} = D = \rho_0(u - v_0)$, а «вытекающей» $I_{\rho_1} = D = \rho_1(u - v_1)$, где u — эйлерова скорость движения разрыва, причем $u - v_0 > 0$ и $u - v_1 > 0$. Первый интеграл (6), как следует из равенства $I_\rho = I_{\rho_0} = I_{\rho_1}$, представляет собой тождество $u = u$.

Обратимся ко второму интегралу (6), записав его для величин с индексами 0 и 1:

$$C_{p_0} = p_0 + \rho_0(u - v_0)^2 - I_\rho u \equiv I_{p_0} - I_\rho u,$$

$$C_{p_1} = p_1 + \rho_1(u - v_1)^2 - I_\rho u \equiv I_{p_1} - I_\rho u.$$

Фигурирующие в этих равенствах величины I_{p_0}, I_{p_1} суть потоки импульса I_p справа и слева от разрыва. Поскольку $C_p = C_{p_0} = C_{p_1}$, то они также равны: $I_p = I_{p_0} = I_{p_1}$. Иначе при переходе через бесконечно тонкую поверхность разрыва частицы вещества получали бы приращение импульса, что возможно лишь при наличии действующих на нем бесконечно больших сил.

Наконец, третий интеграл (6) после несложных выкладок записывается по обе стороны разрыва в виде

$$C_{\varepsilon_0} = I_{\varepsilon_0} - u\left(I_{p_0} + I_{\rho_0}\frac{u}{2}\right), \quad C_{\varepsilon_1} = I_{\varepsilon_1} - u\left(I_{p_1} + I_{\rho_1}\frac{u}{2}\right),$$

где $I_\varepsilon = \rho(u - v)\varepsilon + \rho(u - v)(u - v^2)/2 + p(u - v)$ — поток энергии. Поскольку в бегущей волне величина C_ε постоянна и, как уже установлено выше, постоянны величины I_ρ, I_p, то также постоянен поток энергии I_ε, что означает равенство $I_{\varepsilon_0} = I_{\varepsilon_1}$. Энергия, «втекающая» в разрыв, равна «вытекающей» из него энергии (противное означало бы существование внутри него источников энергии бесконечной интенсивности).

Учитывая результаты, следующие из анализа интегралов (6), приходим к выводу о непрерывности потоков массы, импульса и энергии при переходе через поверхность разрыва газодинамических параметров:

$$I_{\rho_0} = I_{\rho_1}, \quad I_{p_0} = I_{p_1}, \quad I_{\varepsilon_0} = I_{\varepsilon_1}.$$

Заметим, что этот вывод справедлив также и в случае $D = 0$ (*контактный разрыв*).

Для движущегося по массе газа разрыва (*ударной волны*) из установленных равенств нетрудно найти при известной скорости D *условия Гюгонио* — однозначную связь между величинами до и после скачка (см. упр. 2, 3). Эти условия можно также получить и с помощью непосредственного вычисления потоков по обе стороны ударной волны. Однако автомодельные решения типа бегущей волны в силу их свойств позволяют автоматически получать условия Гюгонио для большого числа моделей сплошных сред, в частности для ситуаций, когда скачок происходит не в бесконечно тонком слое, а имеет пространственную протяженность, обусловленную всегда присутствующими в веществе диссипативными процессами. При этом уравнения для бегущей волны описывают еще и структуру переходного слоя.

Заметим, что условия Гюгонио допускают как скачки сжатия, когда давление, плотность и внутренняя энергия газа за ударной волной возрастают, так, формально, и скачки разрежения. Последние, однако, могут реализовываться лишь при наличии в среде соответствующих физических процессов, например, специального типа химических реакций.

Самоподобность, или автомодельность, построенного решения хорошо видна на рис. 68: оно без изменений воспроизводится в различные моменты времени на различных участках вещества. Решение понимается в обобщенном смысле, поскольку оно удовлетворяет дифференциальным уравнениям (5) лишь в областях непрерывного течения. Рассматриваемое в области $-\infty < m < \infty$, оно представляет собой одно из *обобщенных решений* задачи Коши для уравнений газовой динамики. Если же в какой-либо точке с фиксированной массовой координатой, скажем, в точке $m = 0$, задаются соответствующие граничные условия, то его можно трактовать как решение в области $m > 0$ задачи о поршне, вдвигающемся с постоянной скоростью в среду с постоянными параметрами.

Изучим теперь степенное автомодельное решение для частного случая уравнения (4)

$$\frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(k_0 T^{\sigma} \frac{\partial T}{\partial x}\right), \tag{7}$$

описывающего распространение тепла в неограниченной среде ($-\infty < x < \infty$) от мгновенного точечного источника. В момент $t = 0$ в точке $x = 0$ выделяется количество тепла Q_0. Данная постановка полностью совпадает с постановкой аналогичной задачи из п. 5 § 2 гл. II, с той

разницей, что коэффициент теплопроводности (точнее, *температуропроводности*) $k(T) = k_0 T^{\sigma}$, $\sigma > 0$, не постоянен, а является растущей степенной функцией температуры.

Для нахождения автомодельного решения используем П-теорему. Очевидно, оно зависит от четырех определяющих параметров x, t, k_0, Q_0, т. е. $T = T(x, t, k_0, Q_0)$. Данное равенство связывает пять размерных величин, три из которых имеют независимые размерности. Следовательно, согласно П-теореме оно сводится к функциональной зависимости $\Pi = F(\Pi_1)$ между двумя безразмерными величинами, Π, Π_1.

Выберем в качестве размерно зависимых величины Q_0, x, t. Тогда из уравнения (7) для k_0 следует формула размерности $[k_0] = [Q_0]^{-\sigma}[x]^{\sigma-2}/[t]$. Формулу размерности $[T] = [Q_0]/[x]$ для T получаем из условия

$$Q_0 = \int_{-\infty}^{\infty} T(x,t)\,dx, \qquad t \geqslant 0,$$

означающего сохранение первоначально выделившейся энергии в неограниченной среде без источников и стоков тепла. Из П-теоремы и формул размерности имеем

$$\Pi = T(x,t)\,x Q_0^{-1}, \quad \Pi_1 = k_0 t x^{-(2+\sigma)} Q_0^{\sigma}$$

и, с учетом связи $\Pi = F(\Pi_1)$,

$$T(x,t)\,x Q_0^{-1} = F(k_0 t x^{-(2+\sigma)} Q_0^{\sigma}).$$

Перепишем последнее равенство в виде, принятом в табл. 3, вводя обозначение $\xi = \Pi_1^{-1/(\sigma+2)}$:

$$T(x,t) = Q_0 x^{-1} F(\Pi_1) = Q_0 x^{-1}\Phi(\xi) = Q_0^{\frac{2}{2+\sigma}} k_0^{-\frac{1}{2+\sigma}} t^{-\frac{1}{2+\sigma}} \xi\Phi(\xi),$$

или, обозначая $\xi\Phi(\xi)$ через $f(\xi)$,

$$T(x,t) = Q_0^{\frac{2}{2+\sigma}} k_0^{-\frac{1}{2+\sigma}} t^{-\frac{1}{2+\sigma}} f(\xi), \qquad \xi = Q_0^{-\frac{\sigma}{2+\sigma}} k_0^{-\frac{1}{2+\sigma}} t^{-\frac{1}{2+\sigma}} x. \tag{8}$$

Безразмерный инвариант ξ называется *автомодельной переменной*, функция $f(\xi)$ — *безразмерной функцией* (представителем) температуры, размерный множитель перед $f(\xi)$ в (8) — масштабным множителем. *Фиксированным автомодельным состоянием* называется состояние, соответствующее фиксированной автомодельной координате $\xi = \xi_0$.

Анализ размерностей дает возможность получить ценную предварительную информацию о процессе, не находя полностью решения задачи. Так, темп изменения со временем решения в точке $\xi = \xi_0$, в частности, в точке $\xi = 0$, определяется лишь масштабным множителем. Поэтому в начале координат $x = 0$ ($\xi = 0$) температура убывает по заранее известному закону

$$T(0,t) \sim t^{-\frac{1}{2+\sigma}} f(0).$$

Также заранее известна скорость роста со временем координаты x для состояния $\xi = \xi_0$:

$$x(\xi_0) \sim t^{\frac{1}{2+\sigma}}\,\xi_0.$$

Если принять $T(0,t)$ за характерную температуру нагретой области, а $x(\xi_0)$ — за ее характерный размер, то их произведение $T(0,t)\cdot x(\xi_0)$ не зависит от времени, что согласуется с условием постоянства тепловой энергии Q_0, содержащейся в среде.

Подставив (8) в (7) и проведя дифференцирование, получим нелинейное обыкновенное дифференциальное уравнение второго порядка относительно $f(\xi)$:

$$\frac{f}{2+\sigma} + \xi\,\frac{f'}{2+\sigma} = -(f^\sigma f')'.$$

Преобразуем его к виду

$$\frac{1}{(2+\sigma)}(\xi f)' + (f^\sigma f')' = 0$$

и один раз проинтегрируем:

$$\frac{1}{(2+\sigma)}\xi f + f^\sigma f' = C_1.$$

Здесь $C_1 = 0$, так как при $\xi = 0$ решение считается ограниченным ($f(0) < \infty$), гладким и, по построению, симметричным, т. е. $f'(0) = 0$. При $C_1 = 0$ переменные в последнем уравнении разделяются и оно легко интегрируется:

$$f(\xi) = \left[(\xi_\Phi - \xi^2)\,\frac{\sigma}{2(2+\sigma)}\right]^{1/\sigma}, \qquad |\xi| \leqslant \xi_\Phi.$$

Поскольку полученная формула при $|\xi| > \xi_\Phi$ дает отрицательные значения, то в этой области $f(\xi)$ полагается равной нулю (решение в точках $|\xi| = \xi_\Phi$ «сшивается» с тривиальным решением). Величина $\xi_\Phi = \xi_\Phi(\sigma) > 0$ определяется из равенства $\int_{-\infty}^{\infty} f(\xi)\,d\xi = \int_{-\xi_\Phi}^{\xi_\Phi} f(\xi)\,d\xi = 1$ — безразмерного аналога условия постоянства энергии в среде

$$\xi_\Phi = \left[\frac{(2+\sigma)^{\sigma+1}\,2^{1-\sigma}}{\sigma\pi^{\sigma/2}}\,\frac{\Gamma^\sigma(1/2+1/3)}{\Gamma^\sigma(1/\sigma)}\right]^{\frac{1}{\sigma+2}},$$

где Γ — гамма-функция.

Используя (8), приходим к окончательному виду решения задачи о мгновенном точечном источнике тепла в нелинейной среде:

$$T(x,t) = \begin{cases} Q_0^{\frac{2}{2+\sigma}}\,k_0^{-\frac{1}{2+\sigma}}\,t^{-\frac{1}{2+\sigma}}\left\{\left[1-\dfrac{x^2}{x_\Phi^2}\right]\dfrac{\xi_\Phi\sigma}{2(2+\sigma)}\right\}^{1/\sigma}, & |x| \leqslant x_\Phi(t), \\ 0, & |x| > x_\Phi(t). \end{cases} \qquad (9)$$

Здесь $x_\Phi(t) = \xi_\Phi(Q_0^\sigma k_0)^{1/(2+\sigma)}t^{1/(2+\sigma)}$.

Решение (9) обобщенное, так как при достаточно больших σ его производные по x и t в точках $|x| = x_\Phi(t)$ не существуют и в них оно не удовлетворяет уравнению (7) в классическом смысле (см. упр. 4). Однако естественное физическое требование непрерывности потока тепла $W(x,t) = -k_0T^\sigma\,\partial T/\partial x$ в точках $|x| = x_\Phi(t)$, как и для решения (21) § 2 гл. II, выполняется (противное означало бы наличие в этих точках источников или стоков энергии бесконечной интенсивности).

Распространение тепла происходит в виде *волны*, охватывающей со временем все новые и новые участки вещества (рис. 69). Она движется по среде с *конечной скоростью* (ср. с решением (20) § 2 гл. II для случая $\sigma = 0$). *Фронт волны* — точка, отделяющая нагретую часть пространства от холодной, — продвигается по закону $x_\Phi(t) \sim t^{\frac{1}{2+\sigma}}$, отвечая одновременно как фиксированному автомодельному состоянию $\xi = \xi_\Phi$, так и фиксированному физическому состоянию $T(x_\Phi(t),t) = 0$.

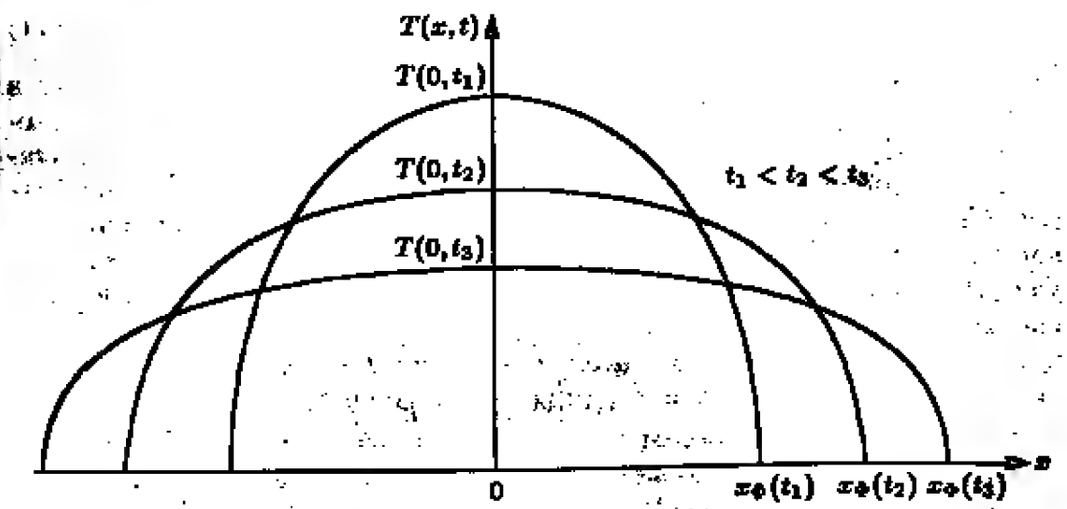


Рис. 69

Самоподобие решения (9) имеет несколько иное геометрическое выражение, чем в предыдущем примере. Кривые на рис. 69 можно совместить в одну (в функцию $f(\xi)$) после их «автомодельной обработки» — соответствующих преобразований растяжения-сжатия как аргументов x, t, так и функции $T(x,t)$.

Особую роль автомодельные решения играют для нелинейных моделей как важные частные решения, раскрывающие те или иные свойства объектов. Их изучение способствует выработке своего рода элементарного языка нелинейных явлений (ударная волна, тепловая волна и т. д.). Однако этим их значение далеко не ограничивается. При определенных условиях они служат *промежуточными асимптотиками* для большого класса формально неавтомодельных процессов. Очевидно, что выделение тепловой энергии в какой-то части вещества не может быть ни мгновенным, ни точечным. Тем не менее ее распространение в среде (с $k(T) = k_0T^\sigma$, $\sigma > 0$), достаточно холодной в момент $t = 0$, по истечении достаточно большого времени и при достаточно больших значениях $x_\Phi(t)$ с хорошей точностью описывает-

ся решением (9). Детали начальной стадии «забываются», и процесс «выходит» на самоподобный режим (на промежуточную асимтотику).

3. Различные режимы распространения возмущений в нелинейных средах. Если нелинейная модель обладает богатыми групповыми свойствами, то появляется возможность получать не отдельные решения, а наборы промежуточных асимптотик с разнообразными свойствами.

Изучим различные режимы распространения тепла, описываемые уравнением (7). Процесс рассматривается в полуограниченной среде $0 < x < \infty$, холодной в момент $t = t_0$:

$$T(x, t_0) = T_0(x) = 0, \qquad 0 < x < \infty, \tag{10}$$

и нагреваемой с левой границы таким образом, что температура в точке $x = 0$ растет со временем по закону

$$T(0, t) = A_0 (t_f - t)^n, \qquad n < 0, \quad t_0 \leqslant t < t_f < \infty. \tag{11}$$

Задача (7), (10), (11) имеет смысл лишь при $t < t_f$, поскольку в конечный момент $t = t_f$ граничная температура обращается в бесконечность. Процессы подобного типа называются *режимами с обострением* и являются математической идеализацией многих реально наблюдаемых явлений, порождаемых геометрией объектов, их нелинейностью и т. д. К ним относятся, например, схождение сферической ударной волны к ее центру, схлопывание пузырьков воздуха в жидкостях и другие виды кумуляции (геометрический фактор). Другой пример: рост численности населения Земли $N(t)$ хорошо описывается в течение нескольких последних веков законом $N(t) = N_0/(t_f - t)$, $t_f = 2026$ (год), следующим из популяционной модели $dN/dt = \alpha_0 N^2$ (см. п. 2 § 6 гл. I; сильная нелинейность).

Решение задачи (7), (10), (11) определяется, очевидно, величинами x, t, k_0, A_0, t_0, t_f. Функциональная связь $T = T(x, t, k_0, A_0, t_0, t_f)$ содержит семь параметров, три из которых являются размерно независимыми. Избавимся от «мешающих» автомодельности параметров t_0, t_f. Возьмем (без ограничения общности) *момент обострения* $t_f = 0$ и будем рассматривать процесс при $-\infty \leqslant t < 0$, т. е. положим $t_0 = -\infty$ (время по-прежнему возрастает, течет от прошлого к будущему). Сформулированная таким образом задача становится, по П-теореме, автомодельной. Проводя выкладки, аналогичные применявшимся при выводе формулы (8), получим вид искомого автомодельного решения:

$$T(x, t) = A_0 (-t)^n f(\xi); \qquad \xi = k_0^{-1/2} A_0^{-\sigma/2} (-t)^{-\frac{1+n\sigma}{2}} x, \quad \xi \geqslant 0. \tag{12}$$

Подстановка (12) в (7) дает для нахождения $f(\xi)$ обыкновенное дифференциальное уравнение второго порядка (с краевыми условиями, вытекающими из (10), (11)):

$$\frac{d}{d\xi}\left(f^\sigma \frac{df}{d\xi}\right) - \frac{1 + n\sigma}{2}\xi \frac{df}{d\xi} + nf = 0, \qquad f(\infty) = 0, \quad f(0) = 1. \tag{13}$$

Уравнение (13) в свою очередь допускает преобразование подобия $\xi' = \alpha\xi$, $f' = \beta f$, $\beta = \alpha^{2/\sigma}$ (растяжение-сжатие) и поэтому сводится к уравнению первого порядка (упр. 5), исследуемого стандартными методами. Анализ показывает, что решение задачи (13) существует при всех $n < 0$, $\sigma > 0$, единственно и монотонно.

Его свойства определяются соотношением между параметрами n (скорость роста граничной температуры) и σ (нелинейность среды). Критическим является значение $n = -1/\sigma$, при котором в уравнении (13) второй член выпадает и оно интегрируется явным образом (упр. 6). Переписывая его с помощью (12) в исходных переменных, имеем

$$T_S(x, t) = \begin{cases} A_0 (-t)^{-1/\sigma} \left(1 - \dfrac{x}{x_\Phi}\right)^{2/\sigma}, & x \leqslant x_\Phi, \\ 0, & x > x_\Phi, \end{cases} \tag{14}$$

где

$$x_\Phi = x_S \equiv \left(2k_0 A_0^\sigma \frac{\sigma + 2}{\sigma}\right)^{1/2}. \tag{15}$$

Первое из краевых условий (13) выполняется «досрочно» (т. е. при $x = x_\Phi(k_0, A_0, \sigma) < \infty$), тепловой поток на фронте равен нулю (так же, как и в решения (9)). Решение (14) обобщенное, поскольку его производные в точке $x = x_\Phi$ при достаточно больших σ не существуют.

При $n < -1/\sigma$ условия на границе между нагретым и холодным веществом также выполняются в конечной точке $\xi_\Phi = \xi_\Phi(n, \sigma) < \infty$ (значение координаты фронта для конкретных n, σ находится численно). В ее окрестности поведение решения описывается асимптотическим выражением, первый член которого — решение упрощенного уравнения, получающегося из (13) в предположении $f(\xi) \to 0$, $\xi \to \xi_\Phi$:

$$f(\xi) = \begin{cases} \left(-\dfrac{1 + n\sigma}{2}\sigma\xi_\Phi\right)^{1/\sigma} (\xi_\Phi - \xi)^{1/\sigma} + \ldots, & \xi \leqslant \xi_\Phi, \\ 0, & \xi > \xi_\Phi. \end{cases} \tag{16}$$

Здесь и далее многоточием обозначены члены более высокого порядка малости. Из (16) видно, что решение при $n < -1/\sigma$ обобщенное с тем же порядком гладкости, что и (9) (гладкость у (14) выше, чем у (9) и (16)). Однако поток тепла в точке фронта, как и для (9) и (14), непрерывен.

В отличие от случая $n \leqslant -1/\sigma$, при $n > -1/\sigma$ условия на фронте волны выполняются лишь при $\xi \to \infty$. Решение — гладкая при всех $\xi \geqslant 0$ функция, его асимптотическое разложение у фронта дается формулой

$$f(\xi) = C\xi^{\frac{2n}{1+n\sigma}} + C_1\xi^{\frac{2n-2}{1+n\sigma}} + \ldots, \qquad 1 + n\sigma > 0, \tag{17}$$

где $C = C(n, \sigma) > 0$ (находится численно) и $C_1 = -C^{\sigma+1} + [2n(2n + n\sigma - 1)]/(1 + n\sigma)^2 < 0$.

Для всех n, σ при $\xi \to 0$ решение имеет асимптотику

$$f(\xi) = 1 + f'(0)\xi + \ldots, \quad f'(0) = \frac{df}{d\xi}(n,\sigma)\bigg|_{\xi=0} < 0,$$

означающую, что поток тепла на границе среды положителен и энергия поступает в вещество.

Перейдем к анализу физических свойств построенных решений, различая случаи $n = -1/\sigma$, $n < -1/\sigma$ и $n > -1/\sigma$.

1) При $n = -1/\sigma$ фронт тепловой волны неподвижен. *Эффективный размер* нагретой области, например, *полуширина*, т. е. точка $x_{эф}(t)$ такая, что $T(x_{эф}(t),t)/T(0,t) = 1/2$, также не изменяется со временем. Автомодельность данного решения, как и всех решений в разделяющихся переменных (имеющих вид $u(x,t) = U(t)\,V(x)$), выражается в одинаковости его пространственных профилей в различные моменты времени. Профили отличаются в данном случае лишь растущей со временем амплитудой $U(t)$. Таким же свойством обладает упорядоченный режим Буссинеска, для которого амплитуда — убывающая функция времени (упр. 2 из § 1 гл. II).

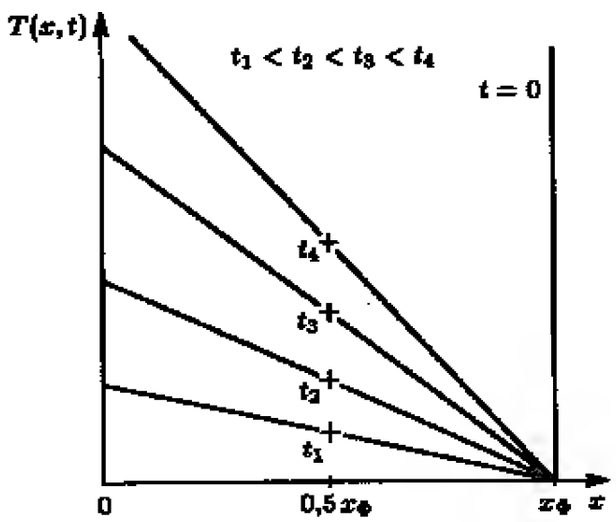


Рис. 70

Решение (14) можно назвать *остановившейся тепловой волной* (см. рис. 70, где показаны профили решения для среды с $\sigma = 2$, крестиками отмечена полуширина). В этом режиме при $t \to 0$ в вещество поступает неограниченное количество энергии; температура и коэффициент теплопроводности при всех $0 \leqslant x < x_S$ стремятся к бесконечности.

Тем не менее тепло не проникает далее точки с координатой $x = x_S$, определяемой интенсивностью граничного режима A_0 и свойствами среды k_0, σ. Решение (14) показывает, что распространение тепла может быть *локализовано* в области конечных размеров — *области локализации* (x_S — *глубина локализации*). Поэтому существует принципиальная возможность концентрировать любое количество энергии в ограниченных участках вещества без распространения ее за пределы зоны локализации.

2) В случае $n < -1/\sigma$ координата фронта волны нагрева неограниченно увеличивается при $t \to 0$:

$$x_ф(t) = k_0^{1/2} A_0^{\sigma/2} (-t)^{\frac{1+n\sigma}{2}} \xi_ф \to \infty.$$

То же самое справедливо для полуширины волны $x_{эф}(t)$ и координат всех точек $x(t,\xi_0)$ с фиксированным автомодельным состоянием $\xi = \xi_0$.

Проследим за изменением решения в фиксированной точке вещества $x = x_0 < \infty$. Отвечающая ей автомодельная координата $\xi(t,x_0)$, как видно из второй формулы (12), стремится к нулю при $t \to 0$. Раскрывая первую формулу (12), получим при $t \to 0$

$$T(x_0,t) = A_0(-t)^n f(\xi(x_0,t)) \to A_0(-t)^n f(0) = A_0(t)^n \to \infty.$$

Таким образом, при приближении к моменту обострения волна охватывает все пространство, температура в любой точке среды неограниченно возрастает, локализация отсутствует. В этом отношении решение аналогично бегущей волне (21) из § 2 гл. II, однако бесконечные параметры достигаются не при $t \to \infty$, а за конечное время. Происходит сверхбыстрый прогрев среды.

3) Совсем по-иному реализуется распространение тепла при $n > -1/\sigma$. Из (12) видно, что полуширина сокращается со временем по закону

$$x_{эф}(t) = k_0^{1/2} A_0^{\sigma/2} (-t)^{\frac{1+n\sigma}{2}} \xi_{эф} \to 0, \qquad t \to 0.$$

Решение представляет собой тепловую волну с уменьшающейся эффективной глубиной прогрева. С течением времени поступающая в вещество энергия сосредоточивается в сокращающейся области вблизи границы (рис. 71). «Фронт» тепловой волны, как следует из (17), (12), находится в точке $x = \infty$ (при $x_ф < \infty$ размер нагретой области со временем уменьшался бы, что в среде без поглощения энергии невозможно).

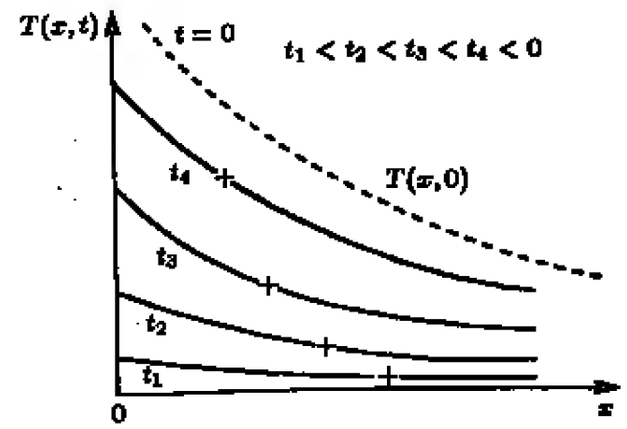


Рис. 71

Для фиксированной точки вещества $x = x_0$ ($0 < x_0 < \infty$) имеем $\xi(x_0,t) \to \infty$, $t \to 0$, т. е. координата $\xi(x_0,t)$ стремится, в отличие от предыдущего случая, к координате фронта. Используя асимптотику (17) и формулу (12), получаем при $t \to 0$

$$T(x_0,t) \to C(k_0^{-n}A_0)^{\frac{1}{1+n\sigma}} x_0^{\frac{2n}{1+n\sigma}} + \\ + C_1(k_0^{1-n}A_0^{1-\sigma})^{\frac{1}{1+n\sigma}} x_0^{\frac{2n+2}{1+n\sigma}}(-t) + \ldots \qquad (18)$$

Несмотря на неограниченный рост температуры в точке $x = 0$, во всей остальной среде она при всех $t \leqslant t_f$ ограничена сверху предельной кривой (первый член в (18); на рис. 71 — штриховая кривая). В случае 3) также осуществляется локализация тепла, но в другом режиме, чем при $n = -1/\sigma$.

Анализ автомодельных решений задачи (7), (10), (11) позволил с помощью простых математических средств установить существование в теплопроводной среде трех принципиально различных режимов

распространения тепла и изучить ряд их важных свойств. Они определяются в основном скоростью подвода энергии к среде при $t \to 0$ и степенью ее нелинейности. Для «медленных» режимов нагрева ($n = -1/\sigma$, *S-режим*; $n > -1/\sigma$, *LS-режим*) имеет место *локализация* тепла, при «быстром» нагреве ($n < -1/\sigma$, *HS-режим*) этот эффект отсутствует.

Наборы разнообразных по своим свойствам автомодельных решений можно получать не только для математических моделей, сводящихся к квазилинейным параболическим уравнениям, но, например, и для уравнений газовой динамики. Изучим обостряющиеся режимы распространения возмущений для простейшего квазилинейного гиперболического уравнения — уравнения Хопфа (32) из § 4 гл. II

$$\frac{\partial p}{\partial t} + k_0 p^\sigma \frac{\partial p}{\partial m} = 0, \tag{19}$$

записанного относительно давления $p = p(m,t)$, $\sigma = (\gamma+1)/(2\gamma) < 1$.

Рассмотрим для него задачу о поршне, сжимающем газ, занимающий полупространство $m > 0$, с нулевым начальным давлением:

$$p(m,t_0) = p_0(m) = 0, \qquad 0 < m < \infty. \tag{20}$$

Поршень расположен в точке $m = 0$, давление на нем растет в режиме с обострением:

$$p(0,t) = A_0\,(t_f - t)^n, \qquad n < 0, \quad t_0 \leqslant t < t_f < \infty. \tag{21}$$

Задача (19)–(21) автомодельна, если подобно предыдущему случаю положить $t_0 = -\infty$ (как и ранее, возьмем без ограничения общности $t_f = 0$). Ее решение в соответствии с П-теоремой имеет вид

$$p(m,t) = A_0\,(-t)^n f(\xi), \qquad \xi = k_0^{-1} A_0^{-\sigma}\,(-t)^{-(1+n\sigma)} m, \quad \xi \geqslant 0. \tag{22}$$

Из (19), (22) получаем для определения $f(\xi)$ уравнение первого порядка

$$\frac{df}{d\xi}\,(f^\sigma + (1+n\sigma)\xi) - nf = 0, \tag{23}$$

которое, будучи записанным в виде

$$\frac{d\xi}{df} = \frac{f^\sigma + (1+n\sigma)\xi}{nf},$$

линейно относительно ξ. Оно имеет общее решение

$$\xi = f^{\sigma + 1/n} - f^\sigma, \tag{24}$$

удовлетворяющее следующему из (21) условию $f(0) = 1$ (при анализе (24) необходимо также следить за выполнением условия $f(\infty) = 0$, вытекающего из (20)). Свойства решений зависят от соотношений между n и σ.

1) При $n = -1/\sigma$ из (24), (22) получаем явное решение в разделяющихся переменных (*остановившаяся волна сжатия*, или газодинамический S-режим):

$$p_S(m,t) = \begin{cases} A_0\,(-t)^{-1/\sigma}\left(1 - \dfrac{m}{m_\Phi}\right)^{-1/\sigma}, & m \leqslant m_\Phi, \\ 0, & m > m_\Phi, \end{cases} \tag{25}$$

где $m_\Phi = m_S \equiv k_0 A_0^\sigma$ — координата фронта волны сжатия, отделяющего пришедший в движение газ от невозмущенного вещества. Фронт волны и ее полуширина постоянны во времени, давление и все остальные газодинамические величины — скорость, плотность и т. д. — неограниченно нарастают в области $0 \leqslant m < m_S$ при $t \to 0$. Однако при $m > m_S$ газ остается неподвижным и холодным при всех $t < 0$ (в эйлеровых координатах это означает, что поршень сжимает прилегающую к нему конечную массу газа m_S в бесконечно тонкий слой, никак не затрагивая остальное вещество). Решение (25) демонстрирует эффект *локализации сжатия* на ограниченной глубине m_S, определяемой параметрами задачи, т. е. эффект, вполне аналогичный локализации тепла в S-режиме.

2) Такая же почти полная аналогия справедлива и в случае $n > -1/\sigma$ (газодинамический LS-режим). Решение, как видно из (24), монотонно убывает с увеличением ξ и обращается в нуль при $\xi = \infty$ (условие $f(\infty) = 0$ выполнено). Фронт волны сжатия расположен в точке $m = \infty$, ее полуширина

$$m_{\text{эф}}(t) = k_0 A_0^\sigma\,(-t)^{1+n\sigma}\xi_{\text{эф}}$$

сокращается до нуля при $t \to 0$ ($\xi_{\text{эф}}$ легко находится из (24) с учетом равенства $f(\xi_{\text{эф}}) = 1/2$). Энергия, сообщаемая газу поршнем, концентрируется в уменьшающейся со временем области вблизи границы.

Для любого $0 < m_0 < \infty$ имеем $\xi(m_0,t) \to \infty$, $t \to 0$. Поэтому, разрешая (24) относительно $f(\xi)$ при $\xi \to \infty$:

$$f(\xi) = \xi^{\frac{1}{1+n\sigma}} + n\,(1+n\sigma)^{-1}\xi^{\frac{n-1}{1+n\sigma}} + \ldots,$$

и переходя с помощью (22) к исходным переменным, получаем при $t \to 0$

$$p(m_0,t) \to \left(\frac{A_0}{k_0^n}\right) m_0^{\frac{n}{1+n\sigma}} + \frac{n}{1+n\sigma}\left(\frac{A_0^{\sigma+1}}{k_0^{n-1}}\right)^{\frac{1}{1+n\sigma}} m_0^{\frac{n-1}{1+n\sigma}}\,(-t) + \ldots$$

Решение стремится снизу к предельной кривой (первый член в полученной формуле), обращаясь в момент $t = 0$ в бесконечность лишь в точке $m = 0$.

3) Для анализа решения при $n < -1/\sigma$ рассмотрим поле интегральных кривых уравнения (23), показанное на рис. 72 (жирной линией выделена кривая I, для которой $f(0) = 1$, штриховая линия — изоклина бесконечности, разделяющая плоскость на области A и B). Из рисунка видно, что непрерывного решения, отвечающего требованию $f(\infty) = 0$,

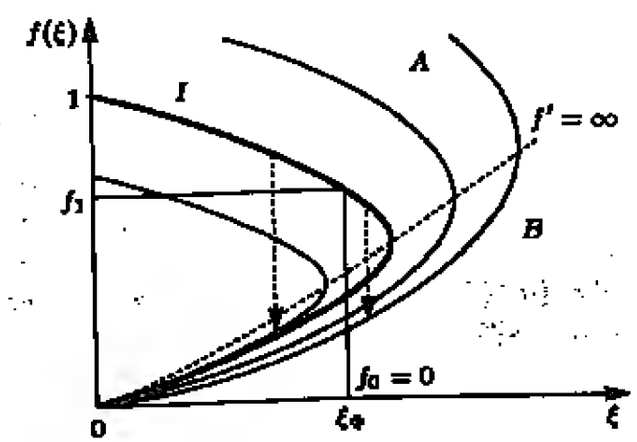


Рис. 72

не существует, так как кривую I нельзя продолжить в область B за линию $f' = \infty$. По этой же причине нельзя построить и разрывные решения, получающиеся с помощью скачков, изображенных стрелками. Таким образом, единственная возможность — переход с кривой I скачком, попадающим на ось абсцисс, и продолжение при $\xi > \xi_\Phi$, ξ_Φ — координата скачка, решения нулем (значение $f(\xi) = 0$ удовлетворяет уравнению (23)).

Условие на скачке для решения уравнения (19) можно получить по аналогии с уравнениями (5), представив (19) в дивергентном виде

$$\frac{\partial p}{\partial t} + \frac{\partial \varphi(p)}{\partial m} = 0,$$

где $\varphi(p) = k_0 p^{\sigma+1}/(\sigma+1)$, и построив для него разрывное решение типа бегущей волны $p(m,t) = p(m - Dt)$. Значения функции p до и после разрыва связаны соотношением

$$D = \frac{\varphi(p_0) - \varphi(p_1)}{p_0 - p_1}.$$

Отсюда для автомодельного решения (22) имеем

$$\bar{D} = -(1+n\sigma)(\sigma+1)\xi_\Phi = \frac{f_0^{\sigma+1} - f_1^{\sigma+1}}{f_0 - f_1},$$

где ξ_Φ — автомодельная координата разрыва, $\bar{D}$ — его обезразмеренная скорость, связанная с мгновенной скоростью $D(t)$ соотношением

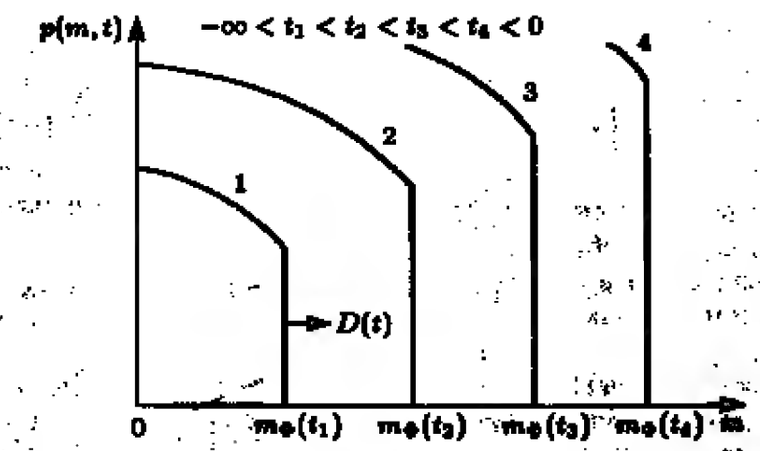


Рис. 73

$D = (\sigma+1)^{-1} k_0 A_0^\sigma (-t)^{n\sigma} \bar{D}$. Для интересующего нас случая $f_0 = 0$ находим связь между f_1 и ξ_Φ

$$f_1 = [-(1+n\sigma)(\sigma+1)\xi_\Phi]^{1/\sigma},$$

используя которую нетрудно установить, что точка (f_1, ξ_Φ) лежит выше линии $f' = \infty$, т. е. искомое разрывное решение действительно существует.

Решение как функция m изображено в различные моменты времени на рис. 73. Его фронт $m_\Phi(t)$ и полуширина $m_{\text{эф}}$ неограниченно растут со временем по закону

$$m_\Phi(t) \sim m_{\text{эф}}(t) \sim (-t)^{1+n\sigma} \to \infty, \qquad t \to 0.$$

Возмущение в пределе охватывает все пространство, решение в любой точке $m_0 < \infty$ стремится к бесконечности при $t \to 0$ (упр. 7).

Таким образом, как и для процессов теплопроводности, в «медленных» режимах с обострением возмущения локализованы, в «быстрых» — локализация отсутствует.

Заметим, что уравнение (19) выводится из уравнений (5) в предположении о непрерывности и изэнтропичности течения (простая волна). Так как построенные решения в случае $n \geqslant -1/\sigma$ непрерывны, то они удовлетворяют (5) и допускают непосредственную газодинамическую интерпретацию (в отличие от разрывного решения при $n < -1/\sigma$).

УПРАЖНЕНИЯ

1. Проверьте прямой подстановкой, что уравнение (4) с $c = c_0$, $\varkappa = \varkappa_0 T^\sigma$ инвариантно к преобразованию растяжения-сжатия при условии $\alpha\beta^{-2}\gamma^\sigma = 1$.

2. Убедитесь, используя (6), что на разрыве, не движущемся по массе газа (контактный разрыв, $D = 0$), скорость и давление непрерывны.

3. Ударная волна, для которой $p_1 \gg p_0$, называется *сильной*. Получите из (6) условие Гюгонио по плотности на таком разрыве (в случае идеального газа $\varepsilon = p/((\gamma-1)\rho)$), полагая для простоты $v_0 = 0$. Покажите, что на сильной ударной волне скачок плотности не зависит от D и дается формулой $\rho_1/\rho_0 = (\gamma+1)/(\gamma-1)$.

4. Установите, что при $\sigma \geqslant 1$ первые производные решения (9) по x и t разрывны в точке $x = x_\Phi(t)$, а при $\sigma \geqslant 1/2$ претерпевают разрыв вторые производные.

5. Проверьте, что замена переменных $\eta = \ln\xi$, $f = \xi^{2/\sigma}\varphi(\eta)$, $\psi = d\varphi/d\eta$ либо замена $\varphi = -\xi f^{-1} f^\sigma df/d\xi$, $\psi = -\xi f^{1-\sigma}/(df/d\xi)$ сводит (13) к уравнению первого порядка вида $d\psi/d\varphi = \psi F(\psi,\varphi)/(\varphi\Phi(\psi,\varphi))$ (во втором случае уравнение исследовать легче, чем в первом, поскольку функции F, Φ содержат простую нелинейность вида $\psi\varphi$).

6. Уравнение (13) при $n = -1/\sigma$ может быть записано в виде $(f^{\sigma+1})'' = -(\sigma+1)nf$. Используя замены $u = f^{\sigma+1}$, $u' = v$, понизьте его порядок и проинтегрируйте полученное уравнение.

7. Убедитесь, используя (22), что при $n < -1/\sigma$ для любой точки $m_0 < \infty$ справедливо $\xi(m_0,t) \to 0$, $t \to 0$ и $p(m_0,t) \to p(0,t) \to \infty$, $t \to 0$.

§ 2. Принцип максимума и теоремы сравнения

Дадим представление о непрерывной зависимости процесса от входных данных. С помощью теорем сравнения и набора автомодельных решений построим замкнутую классификацию режимов с обострением в нелинейных средах. Рассмотрим обобщения автомодельного метода.

1. Формулировка, некоторые следствия. Совокупности промежуточных асимптотик, какими бы разнообразными они ни были, не могут дать описание объекта в общем случае хотя бы потому, что при их получении всегда приходится делать сильные упрощающие предположения (нулевой начальный фон температуры в нагреваемой среде, постоянство течения до и после ударной волны, решение на грани-

це — степенная функция времени и т. д.). Построение достаточно полной картины невозможно без использования *устойчивости* математических моделей, или *непрерывной зависимости решений от входных данных*. Под этим подразумевается свойство решений несильно (в некотором смысле) изменяться при несильном изменении краевых условий, коэффициентов уравнений и других характеристик моделей. Устойчивость, математически выражаемая по-разному для разных ситуаций, является одним из необходимых условий *корректности* моделей, без чего нельзя говорить об их адекватности изучаемому объекту (предполагается, что в определенном смысле устойчив и сам объект). Для устойчивых моделей наборы частных решений служат своего рода ориентирами или границами среди множества всех возможных решений. Это особенно важно, если задача нелинейна и сконструировать ее общее решение из частных нельзя.

Применительно к уравнениям параболического типа это свойство находит свое воплощение в *принципе максимума* и *теоремах сравнения*. Рассмотрим задачу Коши для уравнения нелинейной теплопроводности:

$$\frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(k(T)\frac{\partial T}{\partial x}\right), \qquad k(T) > 0, \quad T > 0;$$
$$-\infty < x < \infty, \quad t > 0, \tag{1}$$
$$T(x,0) - T_0(x) \geqslant 0, \qquad -\infty < x < \infty.$$

Принцип максимума. *Максимум решения $T(x,t)$ (температуры) в любой момент времени не превосходит максимума начальных данных $T_0(x)$ (начальной температуры)*:

$$\max_{t>0,\,-\infty<x<\infty} T(x,t) \leqslant \max_{-\infty<x<\infty} T_0(x). \tag{2}$$

Неравенство (2) имеет очевидный физический смысл. Максимум начального распределения температуры никак не может увеличиться с течением времени, ведь по закону Фурье поток тепла переносит энергию от нагретых участков среды к холодным.

В случае первой краевой задачи для уравнения (1) в полупространстве принцип максимума имеет столь же ясный смысл и означает, что

$$\max_{t>0,\,0<x<\infty} T(x,t) \leqslant \max\left\{\max_{0\leqslant x<\infty} T_0(x),\ \max_{t\geqslant 0} T(0,t)\right\}, \tag{3}$$

где $T_0(x)$ — начальная температура вещества, $T(0,t)$ — задаваемая на его границе температура.

Из принципа максимума следуют теоремы сравнения. Для задачи Коши (2) формулировка выглядит следующим образом.

Пусть $T^{(1)}(x,t)$, $T(x,t)$, $T^{(2)}(x,t)$ — решения задачи Коши, соответствующие начальным данным $T_0^{(1)}(x)$, $T_0(x)$, $T_0^{(2)}(x)$. Тогда если

$$T_0^{(1)}(x) \leqslant T_0(x) \leqslant T_0^{(2)}(x) \quad \textit{для всех } -\infty < x < \infty,$$

то

$$T^{(1)}(x,t) \leqslant T(x,t) \leqslant T^{(2)}(x,t) \quad \textit{для всех } -\infty < x < \infty, \ t > 0. \tag{4}$$

Другими словами, если взять два одинаковых образца теплопроводного материала таких, что начальная температура одного из них не меньше, чем в такой же точке другого, то в любой последующий момент времени это свойство сохранится. Для первой краевой задачи в полупространстве теорема сравнения решений означает следующее.

Выполнение неравенств

$$T_0^{(1)}(x) \leqslant T_0(x) \leqslant T_0^{(2)}(x), \qquad 0 < x < \infty,$$
$$T^{(1)}(0,t) \leqslant T(0,t) \leqslant T^{(2)}(0,t), \qquad t > 0,$$

влечет за собой неравенство

$$T^{(1)}(x,t) \leqslant T(x,t) \leqslant T^{(2)}(x,t) \quad \textit{для всех } 0 \leqslant x < \infty, \ t > 0, \tag{5}$$

где $T^{(1)}(x,t)$, $T(x,t)$, $T^{(2)}(x,t)$ — решения задачи, соответствующие краевым условиям $T^{(1)}(x)$, $T(x)$, $T^{(2)}(x)$ и $T^{(1)}(0,t)$, $T(0,t)$, $T^{(2)}(0,t)$.

Как и принцип максимума, теоремы сравнения имеют непосредственный физический смысл: большее тепловое воздействие на фиксированный объект приводит к формированию в нем большего поля температуры.

Аналоги сформулированных утверждений справедливы и для других задач теории теплопроводности. Используя частные решения каких-либо задач для уравнений параболического и эллиптического типов, можно оценить (ограничить сверху и снизу) решения более общих задач и, не зная их в деталях, сделать выводы общего характера. Непрерывная зависимость решений от входных данных в той или иной форме установлена также для широких классов гиперболических уравнений (например, для уравнения (19) из § 1 она выражается теоремами сравнения, аналогичными теоремам для уравнения (1)).

Из неравенств (4) и свойств решения (9) из § 1 вытекает конечная скорость распространения возмущений в задаче Коши (1) для уравнения (1) с $k(T) = k_0 T^\sigma$, $\sigma > 0$, при условии, что начальное распределение температуры $T_0(x)$ — финитная функция, т. е. $T_0(x) \equiv 0$, $|x| \geqslant R_0 < \infty$. Пусть T_m — максимальное значение функции $T_0(x)$. Тогда, вводя в формулу (9) из § 1 сдвиг по времени на величину t_0 и выбирая постоянные Q_0 и t_0 достаточно большими, нетрудно удовлетворить неравенству $T_m \leqslant \bar{T}(R_0, 0)$, из которого следует неравенство $T_0(x) \leqslant \bar{T}_0(x,0) = \bar{T}_0(x)$, $-\infty < x < \infty$, где через $\bar{T}(x,t)$ обозначено решение задачи о мгновенном точечном источнике тепла. Так как оно мажорирует решение общей задачи Коши (1) по начальным данным, то из теоремы сравнения (4) вытекает $T(x,t) \leqslant \bar{T}(x,t)$, $-\infty < x < \infty$, $t > 0$. Поэтому в любой момент времени $t > 0$ найдется величина $R(t)$ такая, что $T(x,t) \equiv 0$ для всех $|x| \geqslant R(t)$. Это и означает конечную скорость движения фронта тепловой волны.

Путем схожих построений из свойств решения типа бегущей волны (21) § 2 гл. II и неравенств (5) легко устанавливается (при финитной функции $T_0(x)$) конечная скорость распространения тепла в случае первой краевой задачи в полупространстве. Следовательно, этот эффект имеет общий, а не частный характер. Он связан с особенностями уравнения нелинейной теплопроводности (1). Для многих важных процессов температура в некоторых частях вещества может считаться практически равной нулю (например, на начальной стадии сильного взрыва в атмосфере температура вне его зоны ничтожно мала по сравнению с температурой в области, охваченной взрывом). При достаточно сильном росте коэффициента теплопроводности от температуры величина $k(T)$ в этих зонах также фактически равна нулю. Раскрывая правую часть уравнения (1):

$$\frac{\partial T}{\partial t} = k(T)\,\frac{\partial^2 T}{\partial x^2} + k'_T\left(\frac{\partial T}{\partial x}\right)^2,$$

убеждаемся, что в точках x, t, в которых $T(x,t) = 0$ и, таким образом, $k(T(x,t)) = 0$, оно вырождается в уравнение первого порядка (в остальной области (1) — параболическое уравнение второго порядка). Такова математическая подоплека конечной скорости распространения тепла в среду с нулевым температурным фоном. В точках вырождения решение, как уже было показано на частных примерах, следует понимать в обобщенном смысле, в остальной области оно удовлетворяет уравнению (1) в обычном (классическом) смысле.

Более подробно рассмотрим доказательство существования эффекта локализации тепла в задаче Коши. Утверждение состоит в следующем: решение $T(x,t)$ задачи Коши для уравнения

$$\frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(k_0 T^\sigma \frac{\partial T}{\partial x}\right), \qquad \sigma > 0, \quad t > t_0, \quad -\infty < x < \infty, \tag{6}$$

с начальной функцией

$$T(x,t_0) = \begin{cases} T_M\left(1 - \dfrac{x}{x_0}\right)^{2/\sigma}, & |x| \leqslant x_0, \\ 0, & |x| > x_0, \end{cases} \tag{7}$$

локализовано в (первоначальной) области $|x| \leqslant x_0$ в течение времени локализации $t_л$, не меньшего, чем

$$t_л = \frac{x_0^2 \sigma}{2k_0(\sigma+2)T_M^\sigma}, \tag{8}$$

т. е. $T(x,t) \equiv 0$, $|x| > x_0$, $t_0 \leqslant t \leqslant t_0 + t_л$.

В силу (2) максимум $T_m(t)$ решения задачи (6), (7) не превосходит начального при всех $t > t_0$:

$$T_m(t) = T(0,t) \leqslant T_M,$$

причем равенство $T_m(t) = T(0,t)$ следует из симметрии задачи. Рассмотрим решение $T(x,t)$ в области $x > 0$, обозначив его через $T_+(x,t)$. Функцию $T_+(x,t)$ можно, очевидно, трактовать при $t > t_0$ как решение первой краевой задачи для уравнения (6) в области $x > 0$ с начальным условием $T_+(x,t_0)$ из (7) и условием на границе $T_+(0,t)$, удовлетворяющим неравенству

$$T_+(0,t) = T_m(t) = T(0,t) \leqslant T_M.$$

По построению функция $T_+(x,t_0)$ не что иное, как решение (14), (15) из § 1 (автомодельный S-режим $T_S(x,t)$, в котором $A_0 = [x_0^2\sigma/(2k_0(\sigma+2))]^{1/\sigma}$, $x_S = x_0$), взятое в момент $t_0 = -t_л = -x_0^2\sigma/(2k_0(\sigma+2)T_M^\sigma)$.

Сравним решения $T_+(x,t)$ и $T_S(x,t)$ при $t_0 \leqslant t < 0$ (взаимное расположение функций $T(x,t_0)$, $T_+(x,t)$, $T_S(x,t)$, взятых в момент $t > t_0$, показано на рис. 74). Начальные данные в обоих случаях совпадают,

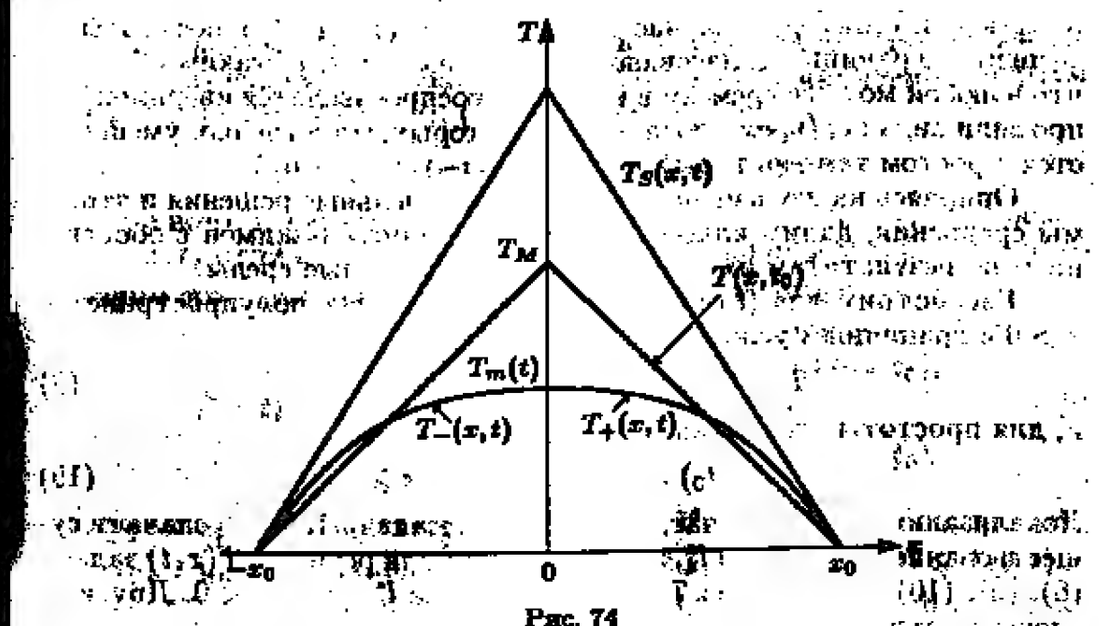


Рис. 74

а граничное условие для $T_+(x,t)$ мажорируется граничным условием для $T_S(x,t)$:

$$T_+(0,t) \leqslant T_M \leqslant A_0(-t)^{-1/\sigma}, \qquad t_0 \leqslant t < 0.$$

Из теоремы сравнения (5) и свойств S-режима имеем

$$T_+(x,t) \equiv 0, \quad x > x_0, \quad t_0 \leqslant t < 0.$$

В силу симметрии $T_-(x,t) = T_+(-x,t)$, где $T_-(x,t)$ — решение $T(x,t)$ в области $x < 0$ (см. рис. 74). С учетом этого окончательно получаем сформулированное выше утверждение

$$T(x,t) \equiv 0, \quad |x| \geqslant x_0, \quad t_0 \leqslant t \leqslant t_0 + t_л.$$

Из теоремы сравнения (4) следует, что любое начальное распределение температуры, мажорируемое функцией (7) (и имеющее совпадающие с ней фронты), также локализовано в области $|x| \leqslant x_0$, а время

локализации оценивается снизу по формуле (8). Таким образом, справедлив вывод общего характера: для среды, теплопередача в которой описывается уравнением (6), всегда можно указать профили температуры, обладающие «инерционностью». Фронты таких профилей, взятых в качестве начальных данных, начинают двигаться по веществу не сразу, а по истечении некоторого времени (см. также упр. 1).

Заметим, что в среде с достаточно сильным поглощением тепла (упр. 2) возможны ситуации, когда размер нагретой области ограничен постоянной величиной при всех $t_0 \leqslant t \leqslant \infty$ (физический смысл этого эффекта очевиден).

2. Классификация режимов с обострением. Инерционность свойственна лишь достаточно «пологим» температурным профилям; это видно из сравнения локализованного распределения (7) и, скажем, решения задачи о мгновенном точечном источнике тепла (формула (9) § 1). Одним из способов формирования подобных профилей может быть воздействие на теплопроводную среду соответствующих граничных режимов. Например, в случае решения (14), (15) из § 1 поступление энергии с границы согласовано со свойствами среды таким образом, что в любой момент времени в веществе воспроизводятся инерционные профили вида (7) (время локализации которых, естественно, уменьшается с ростом температуры).

Опираясь на изученные в п. 3 § 1 автомодельные решения и теоремы сравнения, дадим классификацию граничных режимов с обострением по результатам их воздействия на нелинейные среды.

Рассмотрим для (6) первую краевую задачу в полупространстве $x > 0$ с граничным условием

$$T(0,t) \to \infty, \quad t \to 0, \quad t_0 \leqslant t < 0, \tag{9}$$

и, для простоты, нулевыми начальными данными

$$T(x,t_0) = T_0(x) = 0, \qquad x \geqslant 0. \tag{10}$$

Локализация в этом случае, в отличие от задачи Коши, означает существование постоянной $l^* < \infty$ такой, что для решения $T(x,t)$ задачи (6), (9), (10) справедливо $T(x,t) \equiv 0$ при $l \geqslant l^*$, $t_0 \leqslant t < 0$. Другими словами, тепловые возмущения в течение всего процесса нагрева не проникают далее некоторой конечной глубины l^* (в противном случае локализация отсутствует).

Из теоремы сравнения (5) и свойств автомодельного S-режима вытекает, что если

$$T(0,t) \leqslant A_0(-t)^{-1/\sigma}, \qquad t_0 \leqslant t < 0,$$

то в задаче (6), (9), (10) имеет место локализация тепла на глубине $l^* = x_S$, а ее решение мажорируется функцией $T_S(x,t)$. Если же (9) удовлетворяет неравенству

$$T(0,t) \leqslant A_0(-t)^n, \quad -1/\sigma < n < 0, \quad t_0 \leqslant t \leqslant 0, \tag{11}$$

то решение локализовано на глубине (упр. 3)

$$l^* \leqslant \left(2k_0 A_0 (-t)^{1/\sigma + n} \frac{\sigma+2}{\sigma}\right)^{1/2} \tag{12}$$

и при $0 \leqslant x \leqslant l^*$ для него справедлива оценка (упр. 4)

$$T(x,t) \leqslant C(n,\sigma)(k_0^{-n}A_0)^{\frac{1}{1+n\sigma}} x^{\frac{2n}{1+n\sigma}}, \quad x \leqslant l^*, \quad t_0 \leqslant t < 0, \tag{13}$$

следующая из существования предельной кривой (18) § 1 для автомодельного LS-режима.

Неравенство, предшествующее (11), определяет класс граничных режимов с обострением, приводящих к локализации тепла; неравенство (11) уточняет этот результат: при его выполнении реализуется LS-режим, температура при $t \to 0$ неограниченно растет лишь в точке $x = 0$. При ненулевых, но финитных функциях $T_0(x)$ в (10) оба сделанных выше утверждения о локализации остаются в силе. Необходимо лишь выбрать достаточно большой величину A_0 в мажорирующих решение автомодельных S- и LS-режимах; конкретные оценки (12), (13), естественно, претерпевают изменения. Заметим, что для общего LS-режима фронт волны, в отличие от автомодельного, находится не при $x = \infty$, а в конечной точке.

Эффект локализации в краевых задачах не связан с быстротой нагрева вещества как таковой. Чтобы убедиться в этом, проанализируем поведение решения (6), (9), (10) в случае, когда

$$T(0,t) \geqslant A_0(-t)^n, \quad n < -1/\sigma, \quad t_0 \leqslant t < 0. \tag{14}$$

Прежде всего покажем, что в некоторый момент t_* $(t_0 < t_* < 0)$ решение $T(x,t)$ отлично от нуля в окрестности границы $x = 0$. Для этого сравним его с решением $\bar{T}(x,t)$ типа бегущей волны

$$\bar{T}(x,t) = \begin{cases} \dfrac{D\sigma}{k_0}[D(t-t_0) - x]^{1/\sigma}, & x \leqslant D(t-t_0), \\ 0, & x > D(t-t_0), \end{cases} \tag{15}$$

для уравнения (6), в котором $D = [A_0^\sigma(-t)^{n\sigma-1}k_0/\sigma]^{1/2}$. По построению $\bar{T}(x,t_0) \equiv 0$, а постоянная D подобрана так, чтобы $\bar{T}(0,t) \leqslant T(0,t) \leqslant A_0(-t)^n$, $t_0 \leqslant t < 0$. Следовательно, $T(x,t) \leqslant \bar{T}(x,t)$, $0 \leqslant x < \infty$, $t_0 \leqslant t < 0$ в силу теоремы сравнения (5).

Обратимся теперь к решению (12) из § 1 при $n < -1/\sigma$ (автомодельный HS-режим), обозначив его через $T_a(x,t)$. Выберем в нем величину $A_{0a} \leqslant A_0$ такой, чтобы при $t = t^*$ оно мажорировалось функцией (15), взятой в момент $t = t^*$, т. е. чтобы $T_a(x,t^*) \leqslant \bar{T}(x,t^*)$, $x \geqslant 0$ и, таким образом, $T_a(x,t^*) \leqslant T(x,t^*)$, $x \geqslant 0$. Из (14) имеем $T_a(0,t) \leqslant T(0,t)$ для всех $t > t^*$. Тогда из теоремы сравнения (5) получаем неравенство $T_a(x,t) \leqslant T(x,t)$ для всех $x \geqslant 0$ и $t \geqslant t^*$ (мажорирование при $t \geqslant t^*$ автомодельного HS-режима изучаемым решением). Поскольку $T_a(x,t) \to \infty$, $t \to 0$, $x \geqslant 0$, то и $T(x,t) \to \infty$ при $t \to 0$ в любой точке $x \geqslant 0$.

При выполнении неравенства (14) локализация тепла отсутствует, волна нагрева при $t \to 0$ охватывает все вещество, температура в любой его точке неограниченно растет (данное утверждение тем более справедливо, если $T(x,t_0) \not\equiv 0$).

Этот вывод завершает классификацию граничных режимов с обострением в нелинейных теплопроводных средах. При воздействии на вещество «медленных» S- и LS-режимов ($T(0,t) \leqslant A_0(-t)^{-1/\sigma}$ или $T(0,t) \leqslant A_0(-t)^n$, $-1/\sigma < n < 0$) энергия локализуется в области конечных размеров, в «быстрых» HS-режимах (14) локализация не имеет места.

Для полноты картины поясним, что эффект локализации реализуется и при отказе от требования $T(x,t_0) \equiv 0$, $x \geqslant 0$ (или от требования финитности функции (10)). В этом случае локализацию следует понимать в более общем *эффективном* смысле как существование постоянной $L^* < \infty$ такой, что решение задачи (6), (9) при произвольной ограниченной функции $T_0(x)$ ограничено при $x \geqslant L^*$, $t \geqslant t_0$, несмотря на беспредельный рост решения в точке $x = 0$. Классификация граничных режимов с обострением не зависит от характера начального распределения температуры и остается прежней. В частности, не имеет значения, существует или нет в рассматриваемом процессе конечный фронт тепловой волны.

В качестве примера, иллюстрирующего эти утверждения, рассмотрим для уравнения (6) задачу с граничным условием

$$T(0,t) = A_S(-t)^{-1/\sigma}, \qquad t_0 \leqslant t < 0, \tag{16}$$

отвечающим решению $T_S(x,t)$ — автомодельному S-режиму, но с постоянным начальным фоном температуры

$$T(x,t_0) = T_0 = A_S(-t_0)^{-1/\sigma}. \tag{17}$$

Ее решение $T(x,t)$, очевидно, мажорирует $T_S(x,t)$:

$$T_S(x,t) \leqslant T(x,t), \quad x \geqslant 0, \quad t_0 \leqslant t < 0.$$

Так как температура в точке $x = 0$ для обоих решений одна и та же, то из последнего неравенства вытекает следующее неравенство на производные в этой точке:

$$-\frac{\partial T}{\partial x}\bigg|_{x=0} \leqslant -\frac{\partial T_S}{\partial x}\bigg|_{x=0}, \quad t_0 \leqslant t < 0,$$

а из него — неравенство для потоков тепла на границе

$$W(0,t) = -k(T)\frac{\partial T}{\partial x}\bigg|_{x=0} \leqslant W_S(0,t) = -k(T_S)\frac{\partial T_S}{\partial x}\bigg|_{x=0}, \tag{18}$$
$$t_0 \leqslant t < 0.$$

Физический смысл (18) в том, что при прочих равных условиях первоначально более нагретая среда хуже «воспринимает» поступающую с границы энергию по сравнению с менее нагретой. Проинтегрируем (6) с учетом (18) по x от 0 до ∞ и по t от t_0 до $t < 0$ и (поскольку $W_S(\infty,t) = W(\infty,t) = 0$, $t_0 \leqslant t < 0$) получим

$$\int_0^\infty [T(x,t) - T_0]\,dx \leqslant \int_0^\infty [T_S(x,t) - T_S(x,t_0)]\,dx,$$

или, разбив область интегрирования на части от $x = 0$ до $x = x_S$ и от $x = x_S$ до $x = \infty$ и имея в виду, что $T_S(x,t) \equiv 0$, $x \geqslant x_S$, $t \geqslant t_0$, придем к неравенству

$$\int_0^{x_S} [T(x,t) - T_0]\,dx + \int_{x_S}^\infty [T(x,t) - T_0]\,dx \leqslant \int_0^{x_S} [T_S(x,t) - T_S(x,t_0)]\,dx.$$

Так как $T(x,t) \geqslant T_0$, $x \geqslant 0$, $t \geqslant t_0$ (упр. 5), то данное неравенство можно переписать в виде

$$0 \leqslant \int_{x_S}^\infty [T(x,t) - T_0]\,dx \leqslant \int_0^{x_S} [T_S(x,t) - T(x,t)]\,dx + \int_0^{x_S} [T_0 - T_S(x,t_0)]\,dx,$$

откуда следует, в силу неравенств $T(x,t) \geqslant T_S(x,t)$, $x \geqslant 0$, $t \geqslant t_0$, и $T_0 \geqslant T_S(x,t_0)$, $x \geqslant 0$, ограниченность функции $T(x,t)$ при всех $x \geqslant x_S$, $t_0 \leqslant t \leqslant 0$. Другими словами, в задаче (6), (16), (17) осуществляется *эффективная локализация* тепла на глубине $L^* = x_S$, в точности равной глубине локализации при отсутствии температурного фона.

Классификация режимов с обострением для процессов, описываемых гиперболическим уравнением (19) из § 1, проводится аналогичными методами и приводит к весьма схожим результатам: при действии на среду «медленных» граничных режимов имеет место локализация, в случае «быстрых» режимов этот эффект отсутствует.

3. Расширение «автомодельного метода». Продемонстрированный в пп. 1, 2 подход, основанный на использовании широких классов автомодельных или иных частных решений и непрерывной зависимости процесса от входных данных, допускает разнообразные обобщения.

Покажем, что эффект локализации тепла может осуществляться не только в одномерной, но и в многомерной геометрии, построив соответствующее явное решение уравнения

$$\frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(k_0 T^\sigma \frac{\partial T}{\partial x}\right) + \frac{\partial}{\partial y}\left(k_0 T^\sigma \frac{\partial T}{\partial y}\right) + \frac{\partial}{\partial z}\left(k_0 T^\sigma \frac{\partial T}{\partial z}\right), \tag{19}$$
$$t_0 \leqslant t < 0,$$

рассматриваемого в квадранте $x \geqslant 0$, $y \geqslant 0$, $z \geqslant 0$ (здесь, как и ранее, $T = T(x,y,z,t)$ — температура).

По аналогии с одномерным S-режимом (14), (15) из § 1 будем искать частное решение (19) в разделяющихся переменных, т. е. $T(x,y,z,t) = U(t)f(x,y,z)$. Подстановка этого выражения в (19) дает для $U(t)$ такую же, как и в одномерной геометрии, формулу $U(t) = A_0(-t)^{-1/\sigma}$.

Для функции $f(x,y,z)$ получается сложное эллиптическое уравнение. Поэтому ограничимся простым случаем, когда пространственная часть решения зависит фактически от одного аргумента: $f(x,y,z) = f(\xi)$, $\xi = x + y + z$. Тогда $f(\xi)$ подчиняется уравнению (13) § 1 при

$n = -1/\sigma$, решение которого уже известно. В итоге приходим к многомерному S-режиму

$$T_S(x,y,z,t) = \begin{cases} A_0(-t)^{-1/\sigma}\left(1 - \dfrac{x+y+z}{r_\Phi}\right)^{2/\sigma}, & x+y+z \leqslant r_\Phi, \\ 0, & x+y+z > r_\Phi, \end{cases} \tag{20}$$

где $r_\Phi = r_S \equiv (2k_0A_0^\sigma(\sigma+2)/\sigma)^{1/2}$ вычисляется по формуле для одномерного решения. Решение (20) описывает нагрев трехмерной теплопроводной среды в режиме с обострением, поскольку граничная температура $T(0,y,z,t)$, $T(x,0,z,t)$, $T(x,y,0,t)$ обращается в бесконечность (в области $\xi < r_S$) при $t \to 0$. То же самое происходит в этой области и с решением. Однако во всем остальном пространстве $\xi > r_S$ квадранта $x \geqslant 0$, $y \geqslant 0$, $z \geqslant 0$ температура равна нулю вплоть до момента $t = 0$. Размер области локализации r_S (расстояние от начала координат до плоскости фронта тепловой волны), как и прежде, зависит от свойств среды k_0, σ и интенсивности граничного режима A_0.

Поскольку для решений уравнения (19) справедливы теоремы сравнения, то из (20) сразу определяется класс граничных режимов, приводящих к локализации тепла в рассматриваемой многомерной области. Их дальнейшая классификация вполне аналогична полученным в п. 2 результатам.

Обратимся теперь к более широкой трактовке автомодельных решений, рассмотрев понятие *приближенные автомодельные решения*. Проще всего его ввести на примере задачи о нагреве среды с постоянными теплофизическими свойствами в режиме с обострением:

$$\begin{aligned} &\frac{\partial T}{\partial t} = k_0\frac{\partial^2 T}{\partial x^2}, \qquad 0 < x < \infty, \quad t_0 \leqslant t < 0, \\ &T(0,t) \to \infty, \qquad t \to 0. \end{aligned} \tag{21}$$

Положим без ограничения общности $T(x,t_0) = 0$, $x \geqslant 0$. Общее решение линейной задачи (21) хорошо известно:

$$T(x,t) = \frac{x}{2\sqrt{\pi k_0}}\int\limits_{t_0}^{t} \exp\left(-\frac{x^2}{4k_0(t-\tau)}\right)(t-\tau)^{-3/2}T(0,\tau)\,d\tau. \tag{22}$$

С помощью (22) нетрудно провести классификацию решений (21) в зависимости от вида $T(0,t)$. Видно, что, поскольку при $\tau \to t$, $t \to 0$ первый сомножитель в подынтегральном выражении стремится к нулю, а второй и третий к бесконечности, то возможен любой из изучавшихся в п. 2 режимов распространения тепла.

Так, для $T(0,t) = A_0e^{-a_0/t}$, $a_0 > 0$ — параметр, характеризующий граничный режим, температура неограниченно растет во всех точках $0 \leqslant x \leqslant x_S = 2\sqrt{k_0a_0}$ при $t \to 0$, а в области $x \geqslant x_S$ ограничена для

всех $t \leqslant 0$. Также ограничена тепловая энергия, содержащаяся правее точки $x = x_S$:

$$\lim_{t\to 0}\int\limits_x^\infty T(x',t)\,dx' < \infty, \qquad x > x_S, \quad t \leqslant 0,$$

т. е. осуществляется эффективная локализация тепла в S-режиме с вычисляемой из (22) глубиной $L^* = x_S = 2\sqrt{k_0a_0}$.

Рассмотрев краевое условие более общего вида

$$T(0,t) = A_0\exp(a_0(-t)^n), \qquad n < 0, \tag{23}$$

получим, что при $-1 < n < 0$ реализуется LS-режим, а для $n < -1$ HS-режим.

Анализ интеграла (22) не может дать некоторых важных детальных свойств процесса. Например, для того чтобы получить закон изменения со временем полуширины $x_{\text{эф}}(t)$, надо фактически решить интегральное уравнение (22). Задача (21), (23) не допускает, как легко показать с помощью П-теоремы, и построение автомодельных решений.

Модифицируем (21), (23) следующим образом: вместо (23) возьмем граничное условие в виде $T(0,t) = A_0[\exp(a_0(-t)^n) - 1]$, $n < 0$ (добавление константы сделано для удобства и при $t \to 0$ роли не играет) и проведем замену $V(x,t) = A_0\ln(T(x,t)/A_0 + 1)$. Тогда для $V(x,t)$ получим задачу

$$\begin{aligned} &\frac{\partial V}{\partial t} = k_0\frac{\partial^2 V}{\partial x^2} + \frac{k_0}{A_0}\left(\frac{\partial V}{\partial x}\right)^2, \qquad 0 < x < \infty, \quad t_0 \leqslant t < 0, \\ &V(0,t) = A_0a_0(-t)^n, \qquad n < 0, \quad t_0 \leqslant t < 0, \\ &V(x,t_0) = 0, \qquad 0 \leqslant x < \infty. \end{aligned} \tag{24}$$

Граничное условие (24) — степенная функция времени. Если по аналогии с п. 3 § 1 искать решение задачи (24) в виде, близком к степенному автомодельному решению, то для такого решения при $t \to 0$ первый член в правой части уравнения становится ничтожно малым по сравнению со вторым, и им можно пренебречь (упр. 6).

Более строгий анализ показывает, что действительно решение задачи (24) близко при $t \to 0$ к решениям более простой задачи

$$\begin{aligned} &\frac{\partial u}{\partial t} = \frac{k_0}{A_0}\left(\frac{\partial u}{\partial x}\right)^2, \qquad 0 < x < \infty, \quad -\infty < t < 0, \\ &u(0,t) = A_0a_0(-t)^n, \qquad n < 0, \quad -\infty \leqslant t < 0, \\ &u(x,-\infty) = 0. \end{aligned} \tag{25}$$

В отличие от (24) решение $u(x,t)$ задачи (25) — степенное автомодельное решение, проанализировать которое не представляет труда. Функция $u(x,t)$ является приближенным автомодельным решением для задачи (24) и, следовательно, для исходной задачи (21), (23). Относительная точность описания улучшается с течением времени:

$$\left|1 - \frac{A_0}{T(x,t)}\, e^{u(x,t)/A_0}\right| \to 0, \quad t \to 0, \quad x \geqslant 0.$$

Из свойств $u(x,t)$ вытекает закон для полуширины волны нагрева $x_{эф}(t) = (-t)^{(1-n)/2} \ln 2\,[k_0/(a_0(-n))]^{1/2}$, которая уменьшается во всех трех режимах (ср. с нелинейной средой). Заметим, что построение и анализ приближенных автомодельных решений возможны и используются при изучении широкого класса уравнений параболического типа.

Еще одно важное расширение автомодельного метода основано на *обобщении понятия сравнения решений*. Сущность этого подхода состоит в сравнении решений задач, отвечающих не только разным краевым условиям, как в пп. 1, 2, но и разным уравнениям (в случае уравнения (1) — разным функциям $k(T)$). В определенном смысле речь идет об *устойчивости* процесса теплопередачи по отношению к *возмущениям теплофизических свойств* среды. При этом в качестве одного из сравниваемых решений можно выбрать решение хорошо изученного уравнения (например, (6) или (21)) и получить содержательные результаты для решений более сложных уравнений.

Простейший вариант такого подхода продемонстрируем в случае уравнения (21). Рассмотрим решения $T^{(1)}(x,t)$ и $T^{(2)}(x,t)$ двух краевых задач в полупространстве $x \geqslant 0$:

$$\begin{aligned} &\frac{\partial T^{(1)}}{\partial t} = k_0^{(1)} \frac{\partial^2 T^{(1)}}{\partial x^2}, && x > 0, \quad t > 0, \\ &T^{(1)}(0,t) = T_1^{(1)}(t), && t > 0, \\ &T^{(1)}(x,0) = T_0^{(1)}(x), && x \geqslant 0; \end{aligned} \tag{26}$$

$$\begin{aligned} &\frac{\partial T^{(2)}}{\partial t} = k_0^{(2)} \frac{\partial^2 T^{(2)}}{\partial x^2}, && x > 0, \quad t > 0, \\ &T^{(2)}(0,t) = T_1^{(2)}(t), && t > 0, \\ &T^{(2)}(x,0) = T_0^{(2)}(x), && x \geqslant 0. \end{aligned} \tag{27}$$

Для разности $V(x,t) = T^{(2)}(x,t) - T^{(1)}(x,t)$ из (26), (27) получим соо-

дующую краевую задачу:

$$\begin{aligned} &\frac{\partial V}{\partial t} = k_0^{(1)} \frac{\partial^2 V}{\partial x^2} + (k_0^{(2)} - k_0^{(1)}) \frac{\partial T^{(2)}}{\partial x}, && x > 0, \quad t > 0, \\ &V(0,t) = T_1^{(2)}(t) - T_1^{(1)}(t), && t > 0, \\ &V(x,0) = V_0(x) = T_0^{(2)}(x) - T_0^{(1)}(x), && x \geqslant 0. \end{aligned} \tag{28}$$

Пусть выполнены требования мажорирования решения $T^{(1)}(x,t)$ решением $T^{(2)}(x,t)$ по краевым условиям:

$$\begin{aligned} &T_1^{(1)}(t) \leqslant T_1^{(2)}(t), && t > 0, \\ &T_0^{(1)}(x) \leqslant T_0^{(2)}(x), && x \geqslant 0, \end{aligned} \tag{29}$$

и по коэффициенту теплопроводности:

$$k_0^{(1)} \leqslant k_0^{(2)}. \tag{30}$$

Пусть также для решения $T^{(2)}(x,t)$ справедливо

$$\frac{\partial^2 T^{(2)}}{\partial x^2} \geqslant 0, \quad \frac{\partial T^{(2)}}{\partial t} \geqslant 0 \quad \text{при} \quad x \geqslant 0, \quad t \geqslant 0. \tag{31}$$

Свойство (31) неубывания функции $T^{(2)}(x,t)$ со временем в любой точке $x \geqslant 0$ обеспечивается (при $T_0^{(2)}(x) \equiv 0$) неубыванием по t граничного условия $T_1^{(2)}(t)$. Проверим это, продифференцировав уравнение (27) по t и получив для функции $Z^{(2)}(x,t) = \partial T^{(2)}/\partial t = -W^{(2)}(x,t)/k_0$ задачу

$$\begin{aligned} &\frac{\partial Z^{(2)}}{\partial t} = k_0^{(2)} \frac{\partial^2 Z^{(2)}}{\partial x^2}, && x > 0, \quad t > 0, \\ &Z^{(2)}(0,t) \geqslant 0, && t > 0, \\ &Z^{(2)}(x,0) \equiv 0. \end{aligned} \tag{32}$$

Ее решение, как нетрудно показать (упр. 7), неотрицательно при всех $x \geqslant 0$, $t \geqslant 0$ (в случае $T_0^{(2)} \not\equiv 0$ для выполнения неравенства (31) достаточно наложить еще и условие $\partial^2 T_0(x)/\partial x^2 \geqslant 0$, $x \geqslant 0$).

При выполнении неравенств (29)–(31) задача (28) для $V(x,t)$ становится задачей

$$\begin{aligned} &\frac{\partial V}{\partial t} = k_0^{(1)} \frac{\partial^2 V}{\partial x^2} + f(x,t), && x > 0, \quad t > 0, \\ &V(0,t) \geqslant 0, && t > 0, \\ &V_0(x) \geqslant 0, && x \geqslant 0, \end{aligned} \tag{33}$$

с неотрицательными краевыми условиями и неотрицательной функцией $f(x,t)$ (источниками тепла) в правой части уравнения (33). Ее решение неотрицательно при всех $x \geqslant 0$ и $t \geqslant 0$. Отсюда следует мажорирование решения задачи (26) с коэффициентом $k_0^{(1)}$ решением задачи (27) с коэффициентом $k_0^{(2)}$ во всей рассматриваемой области:

$$T^{(1)}(x,t) \leqslant T^{(2)}(x,t), \qquad t \geqslant 0, \quad t \geqslant 0. \tag{34}$$

В случае уравнения (1) неравенство (34) справедливо при дополнительном к (29)–(31) требовании на коэффициенты $k^{(1)}(T)$, $k^{(2)}(T)$, имеющем вид $[k^{(2)}(T)/k^{(1)}(T)]'_T \geqslant 0$, $T \geqslant 0$.

С помощью подобных теорем сравнения (справедливых и для параболических уравнений общего вида) устанавливается, например, принципиально важный результат: для среды с любыми теплофизическими свойствами всегда можно указать класс граничных режимов, приводящих к локализации тепла, и класс режимов, при действии которых на среду локализация отсутствует. Таким образом, этот эффект носит общий характер.

УПРАЖНЕНИЯ

1. Используя решение (9) из § 1 и теорему сравнения (4), покажите, что локализация возмущений в задаче Коши возможна лишь в течение конечного времени и что при любой начальной функции $T_0(x) \equiv 0$ координата фронта тепловой волны $x_\Phi(t) \to \infty$, $t \to \infty$.

2. В уравнении

$$\frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(k_0 T^\sigma \frac{\partial T}{\partial x}\right) - q_0 T$$

($q_0 > 0$) член $q_0 T$ характеризует зависящую от температуры интенсивность поглощения тепла в среде. Заменами $T = e^{-q_0 t} V$, $\tau = (1 - e^{-\sigma q_0 t})/(\sigma q_0)$ сведите его к уравнению

$$\frac{\partial V}{\partial \tau} = \frac{\partial}{\partial x}\left(k_0 T^\sigma \frac{\partial V}{\partial x}\right).$$

В качестве решения $V(x,\tau)$ уравнения без поглощения возьмите то, которое дается формулой (9) из § 1, и сделайте обратный переход от $V(x,\tau)$ к $T(x,t)$. Убедитесь, что для построенного таким образом решения $T(x,t)$ исходного уравнения $x_\Phi(t) \to \xi_\Phi (Q_0 k_0)^{1/(2+\sigma)} (\sigma q_0)^{-1/(2+\sigma)}$ при $t \to \infty$.

3. Сравнивая краевое условие (11) с решением $T_S(x,t)$ (формулы (14), (15) из § 1) в точке $x = 0$, докажите справедливость оценки (12) при начальных данных (10).

4. Убедитесь, что решение задачи (6), (9), (10) мажорируется в случае неравенства (11) рассмотренным в п. 3 § 1 решением для автомодельного LS-режима.

5. Проверьте неравенство $T(x,t) \geqslant T_0$, $x \geqslant 0$, $t \geqslant t_0$, для решения задачи (6), (16), (17), используя (5) и учитывая, что константа удовлетворяет уравнению (6).

6. Проверьте, что если искать решение уравнения (25) при $n = -1$ в виде $V(x,t) \approx A_0 a_0(t)^{-1} f(x)$ (аналог S-режима), то при $t \to 0$ (24) вырождается в уравнение (25).

7. Докажите от противного неотрицательность решений задач (32) и (33) (используйте теорему сравнения (5) и тот факт, что при нулевых краевых условиях решения — тождественные нули).

§ 3. Метод осреднения

Рассмотрим вариант метода осреднения, применяемый для изучения пространственно-временной динамики локализованных структур. Сформулируем два подхода к осредненному описанию. С их помощью получим классификацию режимов горения в теплопроводных средах.

1. Локализованные структуры в нелинейных средах. Изучавшийся в § 1, 2 эффект локализации возмущений может проявляться не только при внешних воздействиях на среду, заданных в виде соответствующих граничных режимов с обострением, но и благодаря ее собственным нелинейным свойствам. Достаточно сильная нелинейность (см. простой пример в § 6 гл. I) порождает режимы с обострением, которые в свою очередь служат причиной возникновения структур — локализованных в пространстве неоднородностей величин.

Изучим пространственно-временное поведение распределения температуры в теплопроводном веществе с нелинейными источниками тепла. Энергия выделяется в результате горения, протекания химических или иных видов реакций. Среда считается неограниченной (задача Коши), процесс горения одномерным. Он описывается уравнением

$$\frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(k_0 T^\sigma \frac{\partial T}{\partial x}\right) + q_0 T^\beta, \tag{1}$$

$$q_0 > 0, \quad \beta > 1, \quad -\infty < x < \infty, \quad t \leqslant t_0,$$

с начальной функцией

$$T(x,t_0) = T_0(x), \qquad -\infty < x < \infty. \tag{2}$$

Теплофизические характеристики вещества (коэффициент теплопроводности и сильно нелинейный ($\beta > 1$) источник энергии $q_0 T^\beta$) — степенные функции температуры, хорошо аппроксимирующие в ряде случаев реальные зависимости. Горение инициируется начальным распределением температуры $T_0(x) \not\equiv 0$ (в противном случае вещество оставалось бы холодным).

Построим пример *локализованной структуры горения* в задаче (1), (2). Будем искать ее решение в разделяющихся переменных в виде $T(x,t) = U(t)\, f(x)$. Тогда из (1) имеем

$$f \frac{dU}{dt} = U^{\sigma+1} \frac{d}{dx}\left(k_0 f^\sigma \frac{df}{dx}\right) + U^\beta q_0 f^\beta,$$

откуда видно, что переменные разделяются в случае, когда $\beta = \sigma + 1$, и тогда

$$U^{-(\sigma+1)} \frac{dU}{dt} = \frac{1}{f} \frac{d}{dx}\left(k_0 f^\sigma \frac{df}{dx}\right) + q_0 f^\sigma = C, \tag{3}$$

где $C > 0$ (ищутся режимы с растущей во времени температурой). Для $U(t)$ из (3) получаем выражение

$$U(t) = (C_1 - \sigma C t)^{-1/\sigma}, \qquad C_1 > 0, \tag{4}$$

имеющее смысл лишь при $t < C_1/(\sigma C)$ и обращающееся в бесконечность в конечный момент времени $C_1 = \sigma C t$. Таким образом, температура растет в режиме с обострением. В дальнейшем положим без ограничения общности $C = 1/\sigma$ и обозначим $C_1 = t_f$.

Чтобы найти $f(x)$, заменой $f^{\sigma+1} = y$ преобразуем (3) сначала к уравнению

$$\frac{\sigma k_0}{\sigma+1} y'' = y^{\frac{1}{\sigma+1}} - q_0 \sigma y,$$

имеющему вид уравнения колебаний шарика на пружинке с соответствующей внешней силой (см. п. 1 § 4 гл. I), а затем, используя замену $y' = \omega$, к уравнению первого порядка

$$\frac{\sigma k_0}{\sigma+1} \frac{d\omega}{dy} = \frac{y^{\frac{1}{\sigma+1}} - q_0 \sigma y}{\omega}.$$

Его интегрирование дает связь между ω и y:

$$\frac{\sigma k_0}{\sigma+1} \frac{\omega^2}{2} = \frac{\sigma+1}{\sigma+2} y^{\frac{\sigma+2}{\sigma+1}} - q_0 \sigma \frac{y^2}{2} + C_2.$$

Для определения константы C_2 учтем, что функция $\omega = y' = f^\sigma f'(\sigma+1)$ с точностью до числового множителя представляет собой пространственную часть в выражении для потока $W(x,t) = -k_0 T^\sigma \cdot \partial T/\partial x = -k_0 U^{\sigma+1} f^\sigma f'$. На фронте тепловой структуры как температура f, так и тепловой поток должны обращаться в нуль. Следовательно, $\omega = 0$ при $y = 0$ $(f = 0)$ и отсюда $C_2 = 0$. Переходя в последнем уравнении от y и ω обратно к f, получаем (при $C_2 = 0$) квадратуру

$$dx = \pm \frac{(\sigma+1)\sigma k_0\, df}{f\sqrt{2\frac{\sigma+1}{\sigma+2} f^{-\sigma} - q_0 \sigma}},$$

которая интегрируется явным образом:

$$f = A \cos^{2/\sigma} Bx, \quad A = \left[q_0 \frac{\sigma(\sigma+2)}{2(\sigma+1)}\right]^{-1/\sigma}, \quad B = \frac{\sigma}{2}\sqrt{\frac{q_0}{k_0(\sigma+1)}}. \tag{5}$$

Объединяя (4) и (5), приходим к окончательному виду искомого решения

$$T(x,t) = \begin{cases} [q_0(t_f - t)]^{-1/\sigma} \left\{\dfrac{2(\sigma+1)}{\sigma(\sigma+2)} \cos^2 \dfrac{\pi x}{L_T}\right\}^{1/\sigma}, & |x| \leqslant \dfrac{L_T}{2}, \\ 0, & |x| > \dfrac{L_T}{2}, \end{cases} \tag{6}$$

где $L_T = 2\pi\sqrt{k_0/q_0}\sqrt{(\sigma+1)/\sigma^2}$.

Решение (6) описывает немонотонное распределение температуры (структуру) с неподвижным фронтом и неизменной полушириной — автомодельный S-режим горения в нелинейной теплопроводной среде. Несмотря на беспредельный рост температуры в структуре при $t \to t_f$ (при $t \to -\infty$ температура обращается в нуль), процесс горения локализован в конечной области $|x| \leqslant L_T/2$, размер которой определяется параметрами вещества k_0, q_0, σ. Локализация означает отсутствие влияния горящей структуры на участки среды, находящиеся за границами области $|x| \leqslant L_T/2$.

Таким образом, горение нелинейной среды может осуществляться в виде любого числа независимых друг от друга тепловых структур (если их максимумы разделены расстоянием, большим L_T). Причина усложнения процесса, распада его на структуры заключается в *открытости* рассматриваемой системы, в обмене энергией с окружающей средой. В сравнении с такими системами термодинамически замкнутые процессы не допускают возникновения структур. В теплопроводной среде без источников, как видно из свойств решения задачи о мгновенном точечном источнике тепла (п. 2 § 1) и теорем сравнения, распределение температуры при $t \to \infty$ становится пространственно однородным.

По аналогии с автомодельными граничными режимами с обострением из § 1 для среды без источников тепла возможно построение также автомодельных LS- и HS-режимов горения. Они ищутся в виде (упр. 1)

$$T(x,t) = [q_0(t_f - t)]^{-\frac{1}{\beta-1}} f(\xi),$$
$$\xi = \frac{x}{k_0^{1/2} q_0^{m-1} (t_f - t)^m}, \quad m = \frac{\beta - (\sigma+1)}{2(\beta-1)}, \tag{7}$$

и обладают схожими со своими граничными аналогами свойствами: в случае LS-режима $(\beta > \sigma+1)$ полуширина локализованной структуры сокращается со временем, а температура ограничена сверху предельной кривой при всех $t \leqslant t_f$; в случае HS-режима $(\beta < \sigma+1)$ локализация отсутствует.

Но буквальное повторение схемы, применявшейся в § 1, 2, неприемлемо для изучения тепловых структур, по крайней мере в силу двух причин. Существование и свойства автомодельных LS- и HS-режимов горения устанавливаются (из-за сложности соответствующего уравнения; см. упр. 2) гораздо более тонкими, чем в § 1, 2, методами. Кроме того, непосредственное применение автомодельных решений и теорем сравнения для анализа задачи (1), (2) в общем случае не дает завершенных результатов. Например, относительно нетрудно проверить, что в задаче (1), (2) при $\beta = \sigma + 1$ любые начальные возмущения температуры будут развиваться в режиме с обострением (упр. 3). Доказать же их локализацию даже в этом частном случае, используя обычные теоремы сравнения, невозможно.

Действительно, в отличие от граничных режимов с обострением, различные решения задачи (1), (2) могут иметь различные и неизвестные заранее моменты обострения $t_f^{(1)} \neq t_f^{(2)}$ (из (6) видно, что при прочих равных условиях t_f тем меньше, чем больше амплитуда структуры в момент $t = t_0$). Поэтому одно из сравниваемых решений перестает существовать раньше другого, и дальнейшее сравнение теряет смысл (необходимо использовать более сложные методы сравнения).

2. Различные способы осреднения. Для упрощенного анализа пространственно-временных характеристик тепловых структур воспользуемся методом осреднения. Различные варианты этого метода основаны на отказе от точного описания поведения решения как по пространству, так и по времени, и на переходе к некоторым средним характерным величинам, вычисляемым из более простых моделей.

Применительно к тепловым структурам в качестве таких величин можно выбрать пары «амплитуда—полуширина» или «амплитуда—положение фронта».

Рассмотрим первую из них. Будем считать начальную функцию $T_0(x)$ финитной, имеющей максимум в точке $x = 0$ и близкой к симметричной функции. Тогда решение $T(x,t)$ задачи (1), (2) также будет почти симметричным. Приближенное решение ищется в похожем на автомодельный (см. (7)) виде

$$T(x,t) = \psi(t)\,\theta(\xi), \qquad \xi = \frac{|x|}{\varphi(t)}, \tag{8}$$

где $\psi(t)$ и $\varphi(t)$ — зависящие от времени искомые амплитуда и полуширина структуры, $\theta(\xi)$ — некоторая фиксированная финитная монотонно убывающая функция, причем $\theta(0) = 1$.

Потребуем, чтобы (8) удовлетворяло интегральным равенствам (законам сохранения)

$$\int\limits_{-\infty}^{\infty} \frac{\partial T}{\partial t}\,dx - \int\limits_{-\infty}^{\infty} \frac{\partial}{\partial x}\left(T^{\sigma}\frac{\partial T}{\partial x}\right)dx - \int\limits_{-\infty}^{\infty} T^{\beta}\,dx = 0,$$

$$\int\limits_{-\infty}^{\infty} \frac{\partial T}{\partial t}\,T\,dx - \int\limits_{-\infty}^{\infty} \frac{\partial}{\partial x}\left(T^{\sigma}\frac{\partial T}{\partial x}\right)T\,dx - \int\limits_{-\infty}^{\infty} T^{\beta+1}\,dx = 0.$$

Первое их них — закон сохранения энергии, второе моментное уравнение следует из интегрирования (1), умноженного на $T(x,t)$ (ср. с п. 4 § 3 гл. III). Для простоты в (1) положено $k_0 = q_0 = 1$, что не ограничивает общности, так как эквивалентно использованию замены $t' = q_0 t$, $x' = x\,(q_0/k_0)^{1/2}$. Интегрированием двух последних равенств по частям с учетом того, что при $x = \pm\infty$ поток тепла $-T^{\sigma}\,\partial T/\partial x$

равен нулю, получим

$$\frac{d}{dt}\int\limits_{-\infty}^{\infty} T\,dx = \int\limits_{-\infty}^{\infty} T^{\beta}\,dx,$$

$$\frac{1}{2}\frac{d}{dt}\int\limits_{-\infty}^{\infty} T^2\,dx = -\int\limits_{-\infty}^{\infty} T^{\sigma}\left(\frac{\partial T}{\partial x}\right)^2 dx + \int\limits_{-\infty}^{\infty} T^{\beta+1}\,dx. \tag{9}$$

Подстановка (8) в (9) дает систему для $\psi(t)$, $\varphi(t)$:

$$\frac{d}{dt}[\psi(t)\,\varphi(t)] = \nu_1\psi^{\beta}\varphi(t),$$

$$\frac{d}{dt}[\psi^2(t)\,\varphi(t)] = -\nu_2\psi^{\sigma+2}(t)\,\varphi^{-2}(t) + \nu_3\psi^{\beta+1}(t)\,\varphi(t), \tag{10}$$

где ν_1, ν_2, ν_3 — положительные постоянные,

$$\nu_1 = \int\limits_{-\infty}^{\infty} \theta^{\beta}\,d\xi,$$

$$\nu_2 = 2\int\limits_{-\infty}^{\infty} \theta^{\sigma}\left|\frac{d\theta}{d\xi}\right|^2 d\xi \Big/ \int\limits_{-\infty}^{\infty} \theta^2\,d\xi, \tag{11}$$

$$\nu_3 = 2\int\limits_{-\infty}^{\infty} \theta^{\beta+1}\,d\xi \Big/ \int\limits_{-\infty}^{\infty} \theta^2\,d\xi.$$

При этом предполагается, что функция $\theta(\xi)$ такова, что выражения для ν_1, ν_2, ν_3 имеют смысл.

Систему (10) нетрудно разрешить относительно производных:

$$\frac{d\psi}{dt} = \frac{\psi^{\sigma+1}}{\varphi^2}\left[(\nu_3 - \nu_1)\,\psi^{\beta-(\sigma+1)}\varphi^2 - \nu_2\right],$$

$$\frac{d\varphi}{dt} = \frac{\psi^{\sigma}}{\varphi}\left[(2\nu_1 - \nu_3)\,\psi^{\beta-(\sigma+1)}\varphi^2 + \nu_2\right], \quad t \geq t_0 = 0, \tag{12}$$

а от (12) — перейти к уравнению

$$\frac{d\psi}{d\varphi} = -\frac{\psi}{\varphi}\,\frac{a\psi^{\beta-(\sigma+1)}\varphi^2 - 1}{b\psi^{\beta-(\sigma+1)}\varphi^2 - 1}, \quad \psi = 0, \quad \varphi > 0, \tag{13}$$

где $a = (\nu_3 - \nu_1)/\nu_2$, $b = (\nu_3 - 2\nu_1)/\nu_2$. Потребуем выполнения условия $\nu_3 > 2\nu_1$, т. е.

$$a > 0, \quad b > 0. \tag{14}$$

необходимого для того, чтобы система (12) допускала режимы с обострением.

Итак, анализ задачи (1), (2) свелся к изучению обыкновенного дифференциального уравнения первого порядка. Такое же сильное упрощение получается для осреднения «амплитуда—положение фронта». Решение ищется в том же виде

$$T(x,t) = \psi(t)\,\theta(\xi), \qquad \xi = \frac{|x|}{g(t)}, \tag{15}$$

где $\psi(t) > 0$ — амплитуда структуры, а $g(t) > 0$ уже не полуширина, а положение ее движущегося фронта. Функция $\theta(\xi)$ выбирается так, что $\theta(\xi) > 0$, $0 < \xi < 1$, и $\theta(\xi) = 0$, $\xi \geqslant 1$, $\theta(0) = 1$, $\theta'(0) = 0$. В качестве первого интегрального уравнения для ψ и g выберем закон сохранения энергии и аналогично (10) получим

$$\frac{d}{dt}\left[\psi(t)\,g(t)\right] = \nu_1 \psi^{\beta} g(t), \qquad t > 0. \tag{16}$$

Для вывода дополнительного уравнения воспользуемся тем, что $g(t)$ — фронт тепловой волны, и поэтому $W(g(t),t) \equiv 0$, $T(g(t),t) \equiv 0$. Дифференцируя второе из этих тождеств по времени:

$$\frac{\partial T}{\partial g}\frac{dg}{dt} + \frac{\partial T}{\partial t} \equiv 0,$$

используя (1) и записывая производные в виде пределов, придем к равенству

$$\frac{dg}{dt}\lim_{|x|\to g^-(t)}\left[\frac{T(g(t),t) - T(x,t)}{g(t) - x}\right] = \lim_{|x|\to g^-(t)}\left[\frac{W(g(t),t) - W(x,t)}{g(t) - x} - T^{\beta}(x,t)\right].$$

Оно упрощается, если учесть, что $g(t)$ — точка фронта и $T(g(t),t) = W(g(t),t) \equiv 0$:

$$\frac{dg}{dt}\lim_{|x|\to g^-(t)}\frac{T(x,t)}{g(t) - x} = \lim_{|x|\to g^-(t)}\left[\frac{W(x,t)}{g(t) - x} + T^{\beta}(x,t)\right].$$

В окрестности фронта по предположению температура, как и в инертной среде, имеет следующее асимптотическое представление: $T(x,t) \approx (g(t) - |x|)^{1/\sigma}$ (см. в § 1 решения с движущимся фронтом для уравнения (1) без источника). Тогда, как нетрудно проверить, второй член справа в последнем равенстве мал и им можно пренебречь (что и объясняет предположение о малом влиянии источника энергии на структуру решения у фронта). Отсюда получаем формулу

$$\frac{dg}{dt} = \lim_{T\to 0}\frac{W}{T} = -\lim_{T\to 0}\left(T^{\sigma-1}\frac{\partial T}{\partial x}\right) \tag{17}$$

для скорости dg/dt движения фронта волны. Подстановка (15) в (17) дает

$$\frac{dg(t)}{dt} = \nu_4\frac{\psi^{\sigma}(t)}{g(t)}, \qquad t > 0, \tag{18}$$

где $\nu_4 = -(\theta^{\sigma})'(1)/\sigma > 0$.

Разрешая (16) относительно $\psi'(t)$, запишем (16), (18) как

$$\frac{d\psi}{dt} = \nu_1\psi^{\beta} - \nu_4\psi^{\sigma+1}g^{-2}, \qquad \frac{dg}{dt} = \nu_4\psi^{\sigma}g^{-1}, \tag{19}$$

а затем перейдем к уравнению

$$\frac{d\psi}{dg} = \frac{\psi}{g}\left[\mu\psi^{\beta-(\sigma+1)}g^{2} - 1\right], \qquad g > 0, \qquad \mu = \frac{\nu_1}{\nu_4}. \tag{20}$$

Как и в предыдущем случае, эволюция средних характеристик тепловой структуры описывается гораздо более простой моделью, чем исходная.

3. Классификация режимов горения теплопроводной среды. Начнем анализ задачи (1), (2) со случая $\beta = \sigma + 1$, опираясь на первый способ осреднения. Уравнение (13) имеет простой вид:

$$\frac{d\psi}{d\varphi} = -\frac{\psi}{\varphi}\,\frac{a\varphi^2 - 1}{b\varphi^2 - 1}, \qquad \psi > 0, \quad \varphi > 0, \tag{21}$$

и легко интегрируется:

$$C_0 = \psi^{-1}\varphi^{-1}\left[1 - \frac{\nu_3 - 2\nu_1}{\nu_2}\varphi^2\right]^{-\frac{\nu_3}{2(\nu_3 - 2\nu_1)}}$$

где $C_0 \geqslant 0$ — постоянная, определяемая начальными значениями $\varphi(0)$, $\psi(0)$. Характер эволюции тепловой структуры наглядно виден из поведения фазовых траекторий уравнения (21) (рис. 75). Жирной линией показана траектория

$$\varphi \equiv \varphi_S = [\nu_2(\nu_3 - 2\nu_1)]^{1/2} = \frac{L_T}{2}, \tag{22}$$

отвечающая автомодельному S-режиму горения (6) (она же — изоклина бесконечности уравнения (21)), штриховая линия — изоклина нуля $\varphi = a^{-1/2} < \varphi_S$. Справедливость (22) вытекает из того, что в соответствии с (6) функцию $\theta(\xi)$ в (8) естественно взять в виде

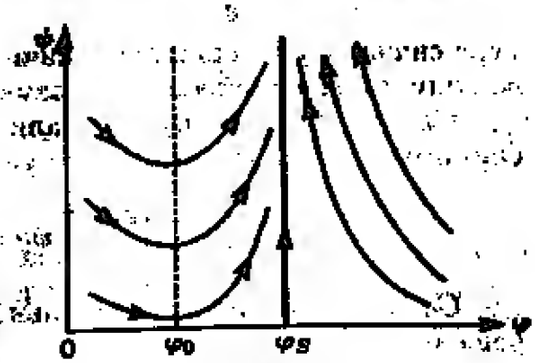


Рис. 75

$$\theta(\xi) = \begin{cases} \cos^{2/\sigma}\dfrac{\pi\xi}{2}, & |\xi| < 1, \\ 0, & |\xi| \geqslant 1. \end{cases}$$

Тогда из (6) получаем для амплитуды и полуширины соотношения

$$\psi(t) = \left[\frac{2(\sigma+1)}{\sigma(\sigma+2)}\right]^{1/\sigma}(t_f - t)^{-1/\sigma}, \quad \varphi(t) = \frac{L_T}{2},$$

а для коэффициентов ν_1, ν_2, ν_3 из (11) —

$$\nu_1 = \frac{\sigma+2}{2(\sigma+1)}, \quad \nu_2 = \frac{\pi^2}{\sigma(\sigma+2)}, \quad \nu_3 = \frac{\sigma+4}{\sigma+2}.$$

Подстановка этих выражений в (21) приводит к тождеству, т. е. автомодельное решение абсолютно точно описывается методом осреднения.

Что же касается неавтомодельных решений, то все они, как видно из рис. 75, развиваются в режиме с обострением, их траектории стремятся к автомодельной: $\varphi(t) \to \varphi_S$, $t \to t_f$. Отсюда следует (упр. 4)

$$\psi(t) \simeq (\sigma\nu_1)^{-1/\sigma}(t_f - t)^{-1/\sigma}, \quad t \to t_f, \tag{23}$$

т. е. на развитой стадии горение протекает по автомодельному закону. На начальной стадии оно может идти более сложным образом: при функциях $T_0(x)$ с полушириной, меньшей φ_0, температура в структуре сначала уменьшается и начинает расти лишь после достижения полушириной *критического* размера $\varphi = \varphi_0$.

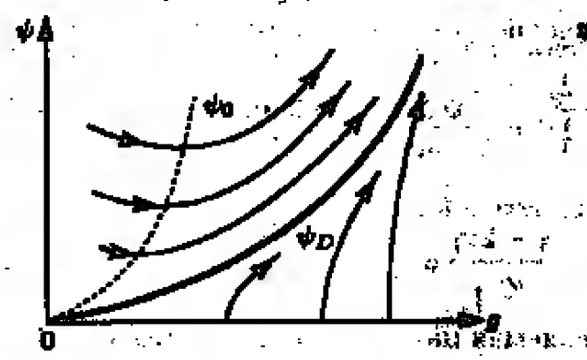


Рис. 76

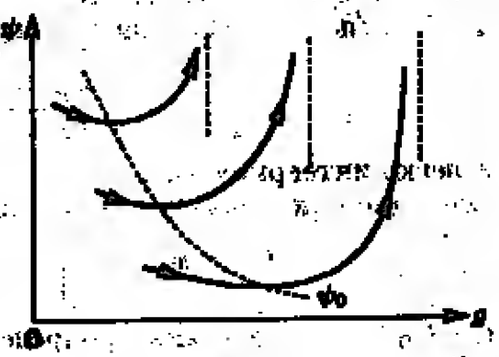


Рис. 77

Случай $\beta \neq \sigma + 1$ будем анализировать с помощью уравнения (20). Оно имеет при $\beta \neq \sigma + 3$ общее решение (упр. 5)

$$\left|\psi^{\beta-(\sigma+1)}g^2 - l_0\right|^{\frac{1}{\sigma+1-\beta}}\psi g = C_0, \quad l_0 = \frac{1}{\mu}\frac{\beta-(\sigma+3)}{\beta-(\sigma+1)}, \tag{24}$$

где $C_0 \geqslant 0$ определяется начальными данными, а при $\beta = \sigma + 3$ — решение

$$\psi^2 = [g^2(C_0 - 2\mu\ln g)]^{-1}, \qquad C_0 = \text{const} > 0. \tag{25}$$

Поведение интегральных кривых уравнения различно в диапазонах $\beta < \sigma + 1$, $\sigma + 1 < \beta < \sigma + 3$, $\beta > \sigma + 3$. На рис. 76, 77, соответствующих случаям $\beta < \sigma + 1$, $\sigma + 1 < \beta < \sigma + 3$, штриховыми линиями показаны изоклины нуля ψ_0, жирной — особая траектория (сепаратриса)

$$\psi = \psi_D = l_0^{\frac{1}{\beta-(\sigma+1)}} g^{-\frac{2}{\beta-(\sigma+1)}}, \tag{26}$$

существующая при $\beta < \sigma + 1$ (и $\beta > \sigma + 3$) и отвечающая значению $C_0 = 0$ в общем решении.

Рассмотрим поведение средних характеристик тепловых структур. В случае $\beta < \sigma + 1$ (рис. 76) все траектории сходятся к сепаратрисе (26), и структура, как следует из (19), асимптотически стремится к автомодельному режиму (см. (7)):

$$\psi(t) \sim (t_f - t)^{\frac{1}{\sigma+1}}, \quad g(t) \sim (t_f - t)^{\frac{\beta-(\sigma+1)}{2(\beta-1)}}, \quad t \to t_f,$$

причем $g(t) \to \infty$ при $t \to t_f$, т. е. локализация тепла в HS-режиме отсутствует.

Более разнообразно поведение тепловых структур LS-режима ($\beta > \sigma + 1$). При $\sigma + 1 < \beta < \sigma + 3$ (рис. 77) любая траектория имеет вертикальную асимптоту с координатой

$$g_* = C_0^{\frac{\beta-(\sigma+1)}{\beta-(\sigma+3)}},$$

т. е. $g(t) \to g_*$, $t \to t_f$, что означает локализацию горения в области $|x| < g_*$. Амплитуда структуры растет по автомодельному закону $\psi \sim (t_f - t)^{-1/(\beta-1)}$, $t \to t_f$ (поведение полуширины при данном способе осреднения, естественно, не описывается).

Вывод о развитии обострившейся локализованной структуры справедлив для части траекторий и в случае $\beta \geqslant \sigma + 3$, причем для $\beta = \sigma + 3$ из (25) получаем $g_* = e^{C_0/(2\mu)}$. Однако при $\beta > \sigma + 3$ (рис. 78; линии ψ_D и ψ_0 имеют тот же смысл, что и на рис. 76) на фазовой плоскости LS-режима существует сепаратриса (26), разделяющая принципиально разные классы решений. Для начальных функций $T_0(x)$, отвечающих области выше линии ψ_D, развивается уже изученный режим с обострением. Если же начальное возмущение имеет либо небольшую амплитуду, либо небольшой размер (область ниже линии ψ_D), то режим с обострением не развивается. Температура в структуре (см. (24)) эволюционирует по закону

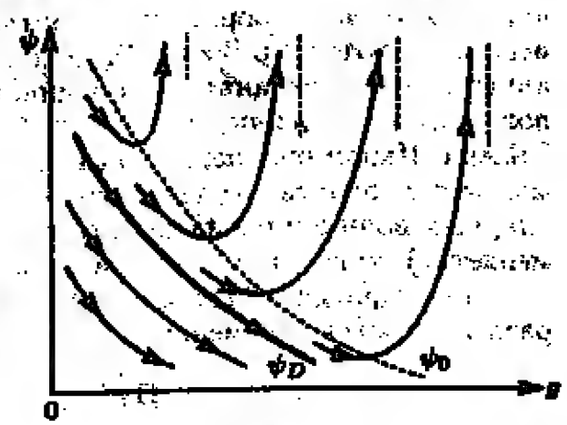


Рис. 78

$$\psi(g) \simeq F_0 g^{-1}, \quad F_0 = C_0 l_0^{-\frac{1}{\sigma+1-\beta}} > 0, \qquad g \to \infty.$$

Отсюда и из (19) получаются асимптотические оценки

$$\psi(t) \sim t^{-\frac{1}{\sigma+2}}, \quad g(t) \sim t^{\frac{1}{\sigma+2}}, \qquad t \to \infty.$$

Подобное поведение характерно для автомодельных решений в среде без выделения энергии, например, для решения задачи о мгновенном точечном источнике тепла из п. 2 § 1. Итак, в случае $\beta > \sigma + 3$ при недостаточно сильном инициировании процесс горения среды затухает.

Заметим, что в HS-режиме и в обостряющемся LS-режиме существует (аналогично S-режиму) критический размер структуры $\varphi = \varphi_0$, при достижении которого ее температура начинает расти.

Метод осреднения позволил получить вполне законченную классификацию структур горения в нелинейной среде. Покажем, что их возникновение не является специфическим свойством процессов, описываемых параболическими уравнениями. Построим простой пример *локализованных газодинамических структур*. Рассмотрим непрерывное (без ударных волн) сжатие конечной массы газа $2M_0$, заключенной внутри цилиндрического поршня (рис. 79). Одномерный процесс описывается системой уравнений

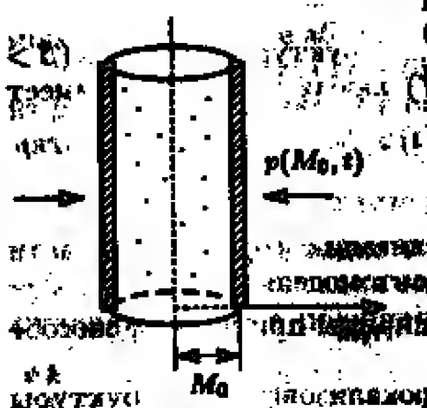


Рис. 79

$$\frac{\partial}{\partial t}\frac{1}{\rho} = \frac{\partial}{\partial m} rv, \quad \frac{\partial v}{\partial t} = -r\frac{\partial p}{\partial m},$$

$$p\rho^{-\gamma} = \varphi(m) \geqslant 0, \tag{27}$$

$$\frac{dr}{dt} = v, \quad 0 < m < M_0, \quad t_0 \leqslant t,$$

где: t — время; m — массовая координата, отсчитываемая от оси симметрии и связанная с эйлеровой координатой соотношением $dm = \rho r\,dr$; ρ, v, p — искомые плотность, скорость и давление газа; $\gamma > 1$ — показатель адиабаты; φ — заданная энтропийная функция. Систему (27) нетрудно получить (аналогично п. 5 § 4 гл. II) из уравнений движения в форме Эйлера. Напомним: первые два уравнения (27) — законы сохранения массы и импульса, третье — интеграл адиабатичности (в отсутствие ударных волн энтропия фиксированной частицы газа со временем не меняется), четвертое — кинематическая связь между r и v.

По аналогии с тепловыми S-режимами построим решение (27) в разделяющихся переменных:

$$p(m,t) = p_1(t)\,\pi(m), \quad v(m,t) = u_1(t)\,u(m),$$

$$r(m,t) = r_1(t)\,R(m), \tag{28}$$

$$0 \leqslant m \leqslant M_0, \quad t_0 \leqslant t < t_f,$$

причем временную часть (28) без существенного ограничения общности (упр. 6) сразу возьмем в виде степенных функций

$$p = p^0(t_f - t)^n \pi(m), \quad v = u^0(t_f - t)^{n_1} u(m),$$

$$r = r^0(t_f - t)^{n_2} R(m), \tag{29}$$

$$0 \leqslant m \leqslant M_0, \quad t_0 \leqslant t < t_f < \infty,$$

и положим в дальнейшем для упрощения выкладок $p^0 = u^0 = r^0 = 1$.

Подставляя эти выражения в (27), получим

$$n = -2, \quad n_1 = \frac{1-\gamma}{\gamma} < 0, \quad n_2 = \frac{1}{\gamma} > 0. \tag{30}$$

В отличие от плоского случая (см. соответствующее решение в п. 3 § 1), показатель n в законе для давления зависит от γ. Для определения π, u, R имеем систему

$$-\frac{2}{\gamma}\varphi(m)^{1/\gamma}\pi^{-1/\gamma} = \frac{d(Ru)}{dm}, \quad \frac{1-\gamma}{\gamma}u = R\frac{d\pi}{dm},$$

$$0 < m \leqslant M_0,$$

которая также решается. Например, пространственный профиль для $\pi(m)$ дается простой формулой

$$\pi(m) = \pi(M_0) + \frac{(\gamma-1)(m-M_0)}{\gamma^2}, \quad 0 \leqslant m \leqslant M_0, \tag{31}$$

где $\pi(M_0)$ определяется из закона для давления на поршне. Заметим, что в силу уравнения $u = -R/\gamma$ скорость газа на оси цилиндра $r = 0$ $(m = 0)$ удовлетворяет естественному условию симметрии $v(0,t) = 0$.

Из (29)–(31) найдем окончательный вид решения для давления:

$$p(m,t) = (t_f - t)^{-2}\left[\pi(M_0) + \frac{(\gamma-1)(m-M_0)}{\gamma^2}\right],$$

$$0 \leqslant m \leqslant M_0, \quad t_0 \leqslant t < t_f,$$

и (с учетом третьего уравнения (27)) для плотности:

$$\rho(m,t) = (t_f - t)^{-2/\gamma}\varphi(m)^{-1/\gamma}\left[\pi(M_0) + \frac{(\gamma-1)(m-M_0)}{\gamma^2}\right]^{1/\gamma},$$

$$0 \leqslant m \leqslant M_0, \quad t_0 \leqslant t < t_f.$$

Подставив в последней формуле для простоты $\pi(M_0) = (\gamma-1)M_0/\gamma^2$ и полагая

$$\rho(m,t) = \left(\frac{\gamma-1}{\gamma^2}\right)^{1/\gamma}(t_f - t)^{-2/\gamma}\left[\frac{m}{\varphi(m)}\right]^{1/\gamma}, \tag{32}$$

$$0 \leqslant m \leqslant M_0, \quad t_0 \leqslant t < t_f.$$

Из (32) видно, что функция $\rho(m,t)$, в отличие от $p(m,t)$, может иметь экстремумы. Степень сжатия участка среды определяется его энтропией $\varphi(m)$ и давлением в нем. Благодаря неизэнтропичности при монотонном профиле давления в волне сжатия возможно достичь большей плотности в областях с меньшим давлением и получить газодинамические структуры (в построенном примере структуры любой

сложности; см., например, упр. 7). Они представляют собой связанные с фиксированной массой газа растущие в режиме с обострением и локализованные немонотонности плотности.

УПРАЖНЕНИЯ

1. Положив формально в задаче (1), (2) $t_0 \to -\infty$ и $T(x,-\infty) \equiv 0$, найдите с помощью П-теоремы представление ее решения в виде (7).

2. Подстановкой (7) в (1) получите для $f(\xi)$ нелинейное обыкновенное дифференциальное уравнение второго порядка. Убедитесь, что оно не допускает преобразования подобия типа растяжение-сжатие (и поэтому не сводится к уравнению первого порядка).

3. Проверьте, что решение задачи (1), (2) при $\beta = \sigma + 1$ и любых $T_0(x) \not\equiv 0$ обращается в бесконечность при $t \to t_f < \infty$ в области, размеры которой не меньше, чем L_T. Для доказательства постройте последовательно ограничивающие его снизу функции — решение задачи о мгновенном точечном источнике тепла и решение (6), а затем примените теорему сравнения.

4. Получите, приближенно решая (12) при $\varphi \to \varphi_S$, $t \to t_f$, формулу (23).

5. Проверьте, что уравнение (20) допускает преобразование подобия растяжение-сжатие. Используя замены вида $\ln\psi = u$, $\ln g = v$, получите его общее решение (24), (25).

6. Подстановкой (28) в (27) проверьте, что функции $p_1(t)$, $v_1(t)$, $r_1(t)$ даются квадратурой, из которой при $t \to t_f < \infty$ асимптотически следуют временные зависимости (29).

7. Подберите примеры такого распределения энтропии $\varphi_N(m) > 0$, $0 < m < M_0$, в (32), чтобы плотность имела любое наперед заданное число экстремумов N. Убедитесь, что при этом функция $\varphi_N(m)$ может быть монотонной.

§ 4. О переходе к дискретным моделям

Дадим представление об элементарных требованиях к численным методам и об основных понятиях теории разностных схем. Рассмотрим некоторые типичные подходы к конструированию дискретных аналогов исходных моделей, используемых для их численного исследования.

1. Необходимость численного моделирования, элементарные понятия теории разностных схем. Как бы глубоки и разнообразны ни были методы качественного анализа математических моделей, область их применимости весьма ограничена. Это — либо простые, главным образом линейные, модели, либо отдельные фрагменты сложных, в том числе нелинейных моделей. Единственным универсальным способом исследования моделей является применение численных методов для нахождения приближенного решения поставленной задачи с помощью средств современной вычислительной техники и информатики.

Доступный «пониманию» компьютера вычислительный алгоритм, т. е. последовательность операций (арифметических, логических и т. д.), в результате выполнения которых находится решение, должен удовлетворять весьма жестким и подчас противоречивым требованиям. К ним относится, прежде всего, необходимость получить решение с заданной точностью за разумное и по возможности минимальное число действий, поскольку время одного расчета должно измеряться минутами и лишь в уникальных случаях — часами. Объемы обрабатываемой при этом информации не могут превышать возможностей емкости

машинной памяти, в процессе вычислений нельзя допускать возникновения не воспринимаемых компьютером слишком больших (малых) чисел, структура алгоритма должна быть достаточно простой и учитывать архитектуру вычислительной системы и т. д.

Только отвечающие этим требованиям вычислительные алгоритмы позволяют проводить всестороннее численное исследование исходной модели, подвергать ее вычислительному эксперименту, проводя ее анализ в самых разных ситуациях и получая исчерпывающую информацию об изучаемом объекте. Такое понимание математического моделирования означает не просто уточнение количественных характеристик явлений, но также изучение основных их качественных свойств. Последнее важно прежде всего для нелинейных объектов, поведение которых может быть весьма разнообразным и неожиданным.

Подчеркнем, что проблемы численного моделирования не снимаются сами собой по мере появления все более мощных и дешевых компьютеров. Это связано, по меньшей мере, с двумя причинами: усложнением выдвигаемых как практикой, так и теорией задач и необходимостью проведения большого числа серий вычислительных экспериментов для достаточно полного изучения объекта.

Поэтому разработка эффективных вычислительных алгоритмов всегда остается одной из ключевых задач математического моделирования. Для их конструирования широко используются методы, идеи и подходы, применяемые при построении исходных математических моделей. Эта связь хорошо прослеживается на примере очень широкого класса моделей — тех, которые сводятся к дифференциальным уравнениям. Для них процесс создания вычислительных алгоритмов состоит из двух главных этапов: на первом строятся *дискретные аналоги* исходных моделей и изучаются их свойства, на втором дискретные уравнения решаются численно (элементарный пример приведен в п. 4 § 6 гл. I).

В дальнейшем основное внимание уделим первому этапу, рассмотрев сначала простейшую краевую задачу для уравнения второго порядка на отрезке

$$\frac{d}{dx}\left(\frac{du}{dx}\right) = f(x), \quad 0 < x < l, \quad u(0) = u_1, \quad u(l) = u_2, \tag{1}$$

предполагая здесь и далее, что ее решение (понимаемое в соответствующем смысле) существует и единственно.

Переход от (1) к дискретной модели разбивается на две стадии. Заменим непрерывную область $0 < x < l$ на дискретную — совокупность конечного числа точек N. Самый простой способ — равномерное деление отрезка $[0,l]$ по правилу $x_i = ih$, $h = l/N$, $0 \leqslant i \leqslant N$. Множество $\bar\omega_h = \omega_h \cup \{x = 0, x = l\}$, $\omega_h = \{x_i\}$, $i \neq 0, N$, этих точек представляет собой (равномерную) *разностную сетку с шагом* h, точки x_i — ее *узлы* (см. рис. 80, на котором черточками отмечены основные узлы сетки, а крестиками — вспомогательные узлы $x_{i+1/2} = (x_{i+1} - x_i)/2$ с полуцелыми индексами). Все фигурирующие в (1) функции рассматри-

ваются теперь как функции не непрерывного аргумента x, а дискретного аргумента x_i (*сеточные функции*), например, аналогом решения

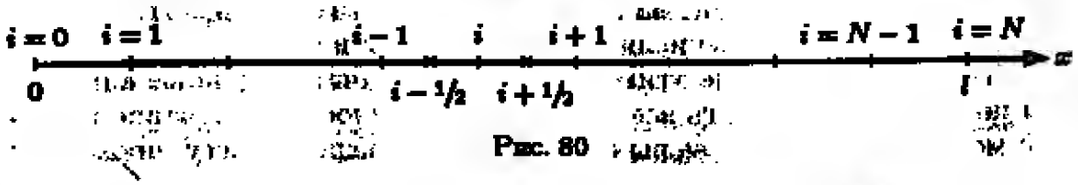


Рис. 80

$u(x)$, $0 \leqslant x \leqslant 1$, служит аппроксимирующее его *приближенное решение* $y(x_i)$, $0 \leqslant i \leqslant N$.

На второй стадии строятся дискретные аналоги дифференциального уравнения (1) и входных данных. Наиболее естественная дискретизация дифференциального оператора — замена производных соответствующими *конечными разностями*. Введем обозначения

$$u_x = \frac{u(x+h)-u(x)}{h}, \quad u_{\bar x} = \frac{u(x)-u(x-h)}{h}, \quad u_{\bar x x} = \frac{u_x - u_{\bar x}}{h}. \qquad (2)$$

Первые два выражения из (2) — дискретная аппроксимация производной du/dx, для получения которой достаточно использовать значения функции $u(x)$ лишь в двух точках (двухточечный *шаблон*). Разлагая $u(x)$ в ряд Тейлора, нетрудно убедиться в том, что

$$\frac{du}{dx}(x) = u_x + O(h), \quad \frac{du}{dx}(x) = u_{\bar x} + O(h),$$

$$\frac{du}{dx}\left(x+\frac{h}{2}\right) = u_x + O(h^2), \quad \frac{du}{dx}\left(x-\frac{h}{2}\right) = u_{\bar x} + O(h^2).$$

Другими словами, du/dx в целых узлах сетки $x = x_i$ аппроксимируется выражениями (2) с первым *порядком аппроксимации*, а в полуцелых точках $x = x_{i+1/2}$, $x = x_{i-1/2}$ (в силу симметрии) — со вторым. Для замены второй производной функции u (третье выражение (2)) требуется, очевидно, трехточечный шаблон $x-h$, x, $x+h$, причем

$$\frac{d}{dx}\left(\frac{du}{dx}\right) = u_{\bar x x} + O(h^2) = \frac{u(x+h)-2u(x)+u(x-h)}{h^2} + O(h^2),$$

т. е. аппроксимация имеет второй порядок (упр. 1).

Дискретная аппроксимация входных данных в рассматриваемом случае не представляет труда и осуществляется точно: $\varphi_i = f(x_i) = f_i$, $i = 0, \ldots, N$; $y_0 = y(0) = u_1$, $y_N = y(l) = u_2$.

Объединяя все эти рассуждения, заменим (1) системой $N-1$ *разностных уравнений* для нахождения $N-1$ неизвестных значений приближенного решения y_i в узлах x_i сетки ω_h:

$$y_{\bar x x} = -\varphi, \quad \text{или} \quad \frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} = -\varphi_i, \qquad (3)$$

$$i = 1, 2, \ldots, N-1, \quad y_0 = u_1, \quad y_N = u_2.$$

Система (3) с краевыми условиями называется *разностной схемой* и служит дискретным аналогом модели (1). Ее решение находится относительно просто (см. п. 2).

Проиллюстрируем, используя (1)–(3), элементарные понятия, связанные с разностными схемами. Сеточная функция $z_i = y_i - u_i$, $i = 0, \ldots, N$, т. е. разница между точным и приближенным решениями в узлах сетки $x = x_i$, называется *погрешностью*. Если $z_i = O(h^\alpha)$, $i = 1, \ldots, N$; $\alpha > 0$, то разностная схема (3) *сходится* (с порядком α), и $z_i \to 0$, $h \to 0$ для всех i: при измельчении сетки y_i сколь угодно хорошо аппроксимирует точное решение $u(x)$ в узлах x_i. Между узлами искомое решение можно при необходимости доопределить с помощью обычной интерполяции. В этом случае построение разностной схемы достигает своей цели.

Чтобы установить сходимость схемы (3), рассмотрим сеточную функцию

$$\psi_i = u_i'' - u_{\bar x x_i}, \qquad i = 1, \ldots, N-1,$$

называемую *погрешностью аппроксимации дифференциального оператора разностным*, или *невязкой* (здесь штрихами обозначено дифференцирование по x, $u_i'' = u''(x_i)$). Если $\psi_i = O(h^\beta)$, $i = 1, \ldots, N-1$; $\beta > 0$, то имеет место аппроксимация: непрерывный оператор приближается дискретным с порядком β (здесь, как уже показано выше, $\beta = 2$), и $\psi_i \to 0$, $h \to 0$, $i = 1, \ldots, N-1$. Функция $f_i - \varphi_i$, $i = 0, \ldots, N$, — погрешность аппроксимации правой части (в изучаемом примере она тождественно равна нулю, как и погрешность аппроксимации краевых условий).

Вычтем во внутренних узлах сетки $i = 1, \ldots, N-1$ уравнение (1) из (3) и получим

$$y_{\bar x x_i} - u_i'' = -\varphi_i + f_i, \quad i = 1, \ldots, N-1,$$

или, учитывая равенства $z = y_i - u_i$, $\psi_i = u_i'' - u_{\bar x x_i}$, $\varphi_i = f_i$, придем к системе разностных уравнений для погрешности z_i с нулевыми краевыми условиями:

$$z_{\bar x x} = \psi, \qquad i = 1, \ldots, N-1, \quad z(0) = z(N) = 0, \qquad (4)$$

правая часть которой — погрешность аппроксимации ψ.

Данное свойство позволяет установить основную связь между уже введенными понятиями сходимости и аппроксимации схемы (3) и понятием ее *устойчивости*. Последнее подразумевает, что для любых допустимых входных данных φ, u_1, u_2 выполняется неравенство

$$\|y\| = \max_i |y_i| \leqslant C\|\varphi\| = C\max_i|\varphi_i|, \qquad i = 0, \ldots, N, \qquad (5)$$

где $C > 0$ — постоянная, не зависящая от i и от h (в данном случае устойчивость по правой части).

Схема (4) — частный случай (3). Поэтому в случае выполнения (5) (см. упр. 2) для нее немедленно имеем

$$|z| \leqslant C|\psi|, \quad \text{или} \quad z_i = O(h^\beta) \to 0, \qquad h \to 0, \quad i = 0, \ldots, N,$$

т. е. из свойств аппроксимации ($\psi = O(h^\beta)$) и устойчивости (5) разностной схемы (3) следует ее сходимость (с тем же порядком, что и порядок аппроксимации, $\alpha = \beta = 2$).

Поясним с помощью качественных рассуждений, что установленная связь распространяется на, вообще говоря, любые классы разностных схем. Пусть некоторая общая (абстрактная) задача

$$\begin{aligned} &Lu = -f(x), \quad x \in G, \quad x = \{x_1, \ldots, x_n\}, \\ &u = \mu(x), \quad x \in \Gamma, \quad \bar{G} = G \cup \Gamma, \end{aligned} \tag{6}$$

в которой L — линейный дифференциальный оператор, действующий в открытой области $x \in G$ (Γ — граница замыкания $\bar{G} = G \cup \Gamma$ области G), f и μ — заданные функции $x \in \mathbb{R}^n$ (правая часть и значение искомого решения на границе), аппроксимируется некоторой разностной схемой

$$\begin{aligned} &L_h y_h = -\varphi_h, \quad x \in \omega_h, \\ &y_h(\gamma_h) = \mu_h, \quad x \in \gamma_h, \quad \bar{\omega}_h = \omega_h \cup \gamma_h, \end{aligned} \tag{7}$$

где $\bar{\omega}_h$, ω_h, γ_h — разностные аналоги для $\bar{G}$, G, Γ, через L_h обозначен соответствующий разностный оператор, а сеточные функции y_h, φ_h, μ_h — аналоги точного решения u и входных данных f, μ.

Введем, как и выше, погрешность $z_h = y_h - u_h$, невязку $\psi_h = (Lu)_h - L_h u_h$ и, считая аппроксимацию правой части и краевых условий точной, получим с учетом (6)

$$L_h z_h = \psi_h, \quad x \in \omega_h; \qquad z_h = 0, \quad x \in \gamma_h. \tag{8}$$

Аналогично случаю схемы (3) правая часть разностной задачи для z представляет собой погрешность аппроксимации ψ. Если схема (7) устойчива, т. е. $\|y_h\|_{(1)} \leqslant C\|\varphi_h\|_{(2)}$ (где символами $\|\cdot\|_{(1)}$ и $\|\cdot\|_{(2)}$ обозначены некоторые, вообще говоря, разные нормы сеточных функций y_h и φ_h), то из (8) следует

$$\|z_h\|_{(1)} \leqslant C\|\psi_h\|_{(2)}.$$

Отсюда видно, что аппроксимация ($\|\psi_h\| \to 0$, $|h| \to 0$) и устойчивость схемы (7) обеспечивают ее сходимость ($\|z_h\| \to 0$, $|h| \to 0$). Такие же утверждения справедливы по отношению и к другим задачам для дифференциальных уравнений, аппроксимируемых разностными схемами, в том числе и к нелинейным (при этом определение устойчивости соответствующим образом модифицируется).

Таким образом, к дискретным моделям предъявляются по меньшей мере два требования — аппроксимация исходной модели и устойчивость. Тогда при достаточно точном численном решении разностных уравнений (как правило, они представляют собой системы линейных и нелинейных алгебраических уравнений порядка N, где N — число узлов сетки) и достаточно малых шагах получается достаточно точное приближенное решение. Подчеркнем: число узлов не может быть слишком большим (а ее шаги слишком маленькими), так как

численное решение необходимо найти за приемлемое число операций, т. е. используя реальные сетки. Рассмотрим некоторые типичные методы построения дискретных моделей.

2. Непосредственная формальная аппроксимация. Этот, исторически первый и, как следует из названия, легко интерпретируемый метод прост, ясен и часто дает дискретные модели хорошего качества. Покажем это, построив разностную аппроксимацию первой краевой задачи на отрезке $[0, l]$ для уравнения теплопроводности

$$\begin{aligned} &\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right), \qquad 0 < x < l, \quad 0 < t \leqslant T, \\ &u(0,t) = u_1(t), \quad u(0,l) = u_2(t), \quad 0 \leqslant t \leqslant T, \\ &u(x,0) = u_0(x), \quad 0 \leqslant x \leqslant l. \end{aligned} \tag{9}$$

Ее решение $u(x,t)$ предполагается существующим и единственным в области $0 < x < l$ при всех $0 < t \leqslant T$ (рис. 81).

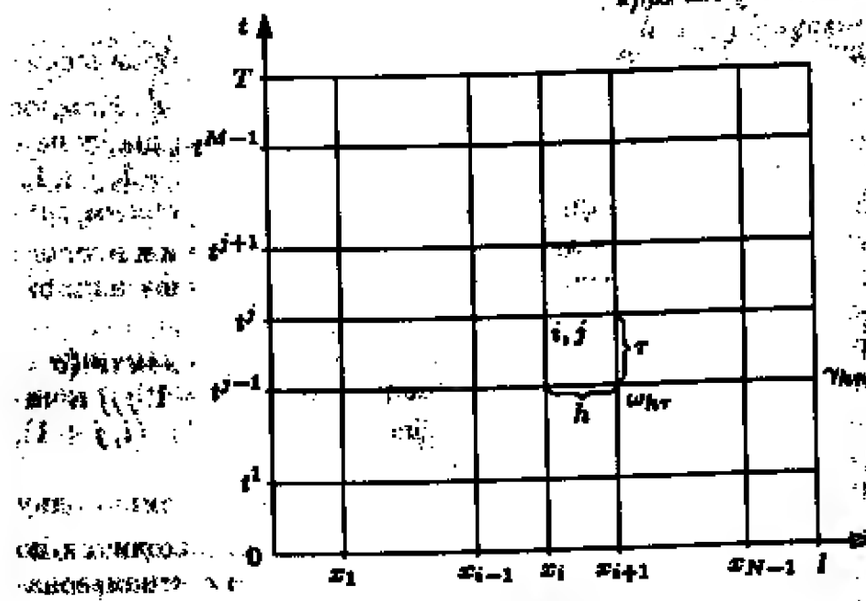


Рис. 81

Выберем простейшую дискретизацию расчетной области, разбив ее равномерно $N-1$ вертикальными и M горизонтальными линиями. Точки их пересечения между собой и с отрезками $[0, l]$, $[0, T]$, $[l, T]$ дадут узлы сетки $\bar{\omega}_{h\tau}$. Такая сетка называется *равномерной с шагом по пространству* $h = l/N$, $x_i = ih$, $0 \leqslant i \leqslant N$, *и по времени* $\tau = T/M$, $t^j = j\tau$, $0 \leqslant j \leqslant M$; совокупность узлов с одинаковым индексом j — *временным слоем* (рис. 81). В граничных (принадлежащих отрезкам $[0, l]$, $[0, T]$, $[l, T]$) узлах функция $u(x,t)$ известна из краевых условий (9), и аппроксимация очевидна: $y_0^j = u_1(t^j)$, $y_N^j = u_2(t^j)$, $j = 0, 1, \ldots, M$; $y_i^0 = u_0(x_i)$, $0 \leqslant i \leqslant N$.

Приближенное решение y_i^j необходимо найти на множестве внутренних узлов $\omega_{h\tau}$ сетки $\bar\omega_{h\tau} = \omega_{h\tau} \cup \gamma_{h\tau}$. Проведем естественную аппроксимацию дифференциального оператора разностным, записывая последний одинаково в любой точке (x_i, t^j) сетки $\omega_{h\tau}$. Производную по времени заменим первой разностью: $\partial u/\partial t \approx (y_i^{j+1} - y_i^j)/\tau$. Учитывая, что в ней фигурируют значения разностного решения с двух временных слоев, вторую производную по x заменим суммой выражений (полученных в п. 1), взятых на $(j+1)$-м и j-м слоях: $\partial^2 u/\partial x^2 \approx \approx \sigma y_{\bar x x}^{j+1} + (1-\sigma) y_{\bar x x}^j$, где $0 \leqslant \sigma \leqslant 1$. В итоге вместо (9) придем к *схеме с весами*

$$\frac{y_i^{j+1} - y_i^j}{\tau} = \sigma \frac{y_{i+1}^{j+1} - 2y_i^{j+1} + y_{i-1}^{j+1}}{h^2} + (1-\sigma)\frac{y_{i+1}^j - 2y_i^j + y_{i-1}^j}{h^2},$$

$$(x_i, t^j) \in \omega_{h\tau}, \tag{10}$$

$$y_0^j = u_1(0, t^j), \quad y_N^j = u_2(l, t^j), \quad j = 0, 1, \ldots, M;$$

$$y_i^0 = u_0(x_i), \quad 0 \leqslant i \leqslant N,$$

представляющей собой систему $(N-1)M$ линейных алгебраических уравнений для нахождения стольких же значений функции y_i^j. Каждое из уравнений (10) записано при $\sigma \neq 0, 1$ на шеститочечном шаблоне с использованием узлов с индексами $(i-1, j+1)$, $(i, j+1)$, $(i+1, j+1)$, $(i-1, j)$, (i, j), $(i+1, j)$. Погрешность ее аппроксимации в общем случае есть величина $O(\tau + h^2)$, т. е. первого порядка по времени и второго по пространству. Для *симметричной схемы* с $\sigma = 1/2$ порядок аппроксимации по времени увеличивается до $O(\tau^2)$ (упр. 3).

При $\sigma = 0, 1$ из (10) получаются более простые чисто *явная* ($\sigma = 0$, четырехточечный шаблон $(i, j+1)$, $(i-1, j)$, (i, j), $(i+1, j)$) и чисто *неявная* ($\sigma = 1$, четырехточечный шаблон $(i-1, j+1)$, $(i, j+1)$, $(i+1, j+1)$, (i, j)) схемы.

При $\sigma = 0$ каждое из уравнений схемы (10) содержит лишь одну неизвестную величину y_i^{j+1}. Поэтому ее решение легко находится по явным формулам при переходе от j-го к $(j+1)$-му слою с использованием известных значений решения на границе (на слое $j = 0$ решение известно из начальных данных).

В случае неявной схемы (и всех схем с $\sigma \neq 0$) разностные уравнения содержат три неизвестных величины: y_{i-1}^{j+1}, y_i^{j+1}, y_{i+1}^{j+1}; на каждом временном слое получается задача типа (3). Подобные задачи для трехточечных уравнений вида

$$A_i y_{i-1} - C_i y_i + B_i y_{i+1} = -F_i, \qquad i = 1, \ldots, N-1,$$
$$y_0 = \varkappa_1 y_1 + \mu_1, \quad y_N = \varkappa_2 y_{N-1} + \mu_2, \tag{11}$$

где $A_i \neq 0$, $B_i \neq 0$, $i = 1, \ldots, N-1$, при условиях $|C_i| \geqslant |A_i| + |B_i|$, $i = 1, \ldots, N-1$; $|\varkappa_\alpha| \leqslant 1$, $\alpha = 1, 2$; $|\varkappa_1| + |\varkappa_2| < 2$ (для (3) и (10) эти

условия выполнены) относительно легко решаются методом прогонки. У решения (11) предполагается наличие рекуррентной зависимости вида

$$y_i = \alpha_{i+1} y_{i+1} + \beta_{i+1}, \tag{12}$$

подстановка которой в (11) дает для коэффициентов α_{i+1} и β_{i+1} рекуррентные соотношения

$$\alpha_{i+1} = \frac{B_i}{C_i - \alpha_i A_i}, \quad \beta_{i+1} = \frac{A_i \beta_i + F_i}{C_i - \alpha_i A_i}, \quad i = 1, \ldots, N_1.$$

Из них с помощью краевого условия при $i = 0$ находятся α_{i+1}, β_{i+1} во всех узлах сетки (прямая прогонка). Далее, с помощью условия в точке $i = N$ при известных α_{i+1}, β_{i+1} по формуле (12) вычисляются значения $y_N, y_{N-1}, \ldots, y_1, y_0$ (обратная прогонка). Заметим: простота решения алгебраической системы (10) обусловлена простой (трехдиагональной) структурой ее матрицы.

Несколько более сложными, чем в п. 1, рассуждениями устанавливается устойчивость схемы (10), причем существует качественная разница между случаями $\sigma = 0$ и $\sigma = 1$. Чисто неявная схема устойчива при любом соотношении между шагами h и τ (*безусловная устойчивость*), в то время как для чисто явной схемы необходимо выполнение неравенства $\tau \leqslant Ch^2$, $C > 0$ — некоторая постоянная (*условная устойчивость*). Данное требование (типичное для явных разностных схем, порождаемых параболическими уравнениями) может накладывать слишком жесткие ограничения на шаг по времени. Поэтому, несмотря на свою простоту, явные схемы для решения этих задач практически не используются.

Аналогичным образом строятся разностные схемы типа (10) для отличных от (9) краевых задач и более общих параболических уравнений. Например, для уравнения нелинейной теплопроводности

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(k(u)\frac{\partial u}{\partial x}\right)$$

одна из очевидных аппроксимаций первой краевой задачи выглядит следующим образом:

$$\frac{y_i^{j+1} - y_i^j}{\tau} = \frac{1}{h}\left[k_{i+1/2}^j \frac{y_{i+1}^{j+1} - y_i^{j+1}}{h} - k_{i-1/2}^j \frac{y_i^{j+1} - y_{i-1}^{j+1}}{h}\right],$$

$$(x_i, t^j) \in \omega_{h\tau},$$

$$y_0^j = u_1(0, t^j), \quad y_N^j = u_2(l, t^j), \quad j = 0, 1, \ldots, M;$$

$$y_i^0 = u_0(x_i), \quad 0 \leqslant i \leqslant N,$$

где $k_{i+1/2}^j$ и $k_{i-1/2}^j$ — некоторые приближения коэффициента теплопроводности в полуцелых узлах $i + 1/2$, $i - 1/2$, взятые для простоты на

j-м слое. Данная неявная устойчивая схема легко решается, как и (10), методом прогонки. Если же коэффициент теплопроводности аппроксимируется на $(j+1)$-м слое, то она становится нелинейной и решается с помощью соответствующих методов *последовательных приближений* (*итерационных процедур*).

Итак, на основе прямой аппроксимации в ряде случаев просто получаются дискретные модели, обладающие нужными качествами. В то же время его формальное применение может привести также и к построению дискретных аналогов исходных моделей, не имеющих ничего общего со своими прообразами. Возьмем весьма естественную аппроксимацию уравнения (9) — разностное уравнение

$$\frac{y_i^{j+1}-y_i^{j-1}}{2\tau}=\frac{y_{i+1}^j-2y_i^j+y_{i-1}^j}{h^2}, \tag{13}$$

записанное на пятиточечном шаблоне $(i,j+1)$, $(i+1,j)$, (i,j), $(i-1,j)$, $(i,j-1)$.

Вместе с соответствующими краевыми условиями оно образует *трехслойную схему* (в отличие от *двухслойной схемы* (10)), имеющую погрешность аппроксимации $O(\tau^2+h^2)$ и легко решаемую по явным формулам. Однако данная схема непригодна, так как неустойчива при любых h и τ (*абсолютно неустойчива*). Ее решение при ограниченных краевых условиях может становиться сколь угодно большим, когда $j\to\infty$ $(t\to\infty)$, что противоречит принципу максимума для уравнения (9).

Поясним это, представляя решение уравнения (13) с нулевыми граничными условиями как сумму частных решений (гармоник), каждая из которых имеет вид $y_{(k)}(x,t)=T_{(k)}(t)\,X^{(k)}(x)$, $k=1,2,\ldots,N-1$ (разделение переменных). Тогда, подставляя $y_{(k)}(x,t)$ в (13) и разделяя переменные, для любой гармоники имеем

$$\frac{T_{(k)}^{j+1}-T_{(k)}^{j-1}}{2\tau T_{(k)}^j}=\frac{X_{i+1}^{(k)}-2X_i^{(k)}+X_{i-1}^{(k)}}{h^2X_i^{(k)}}=-\lambda_k, \tag{14}$$

где $\lambda_k=4/(h^2)\cdot\sin(\pi kh/2)>0$ — параметр разделения, или собственное значение для гармоники с номером k (упр. 4).

Временная часть решения $y_{(k)}(x,t)$, как следует из (14), подчиняется уравнению

$$T_{(k)}^{j+1}-T_{(k)}^j=-\alpha_kT_{(k)}^j,\qquad \alpha_k=2\tau\lambda_k>0,$$

имеющему частное решение вида $T_{(k)}^{j+1}=q_kT_{(k)}^j$ (из данной связи следует $T_{(k)}^{j+1}=q_k^{j+1}T_{(k)}^0$). При этом q_k должна удовлетворять уравнению $q_k^2+\alpha_kq_k-1=0$ с действительными корнями, один из которых по модулю больше единицы при любых α_k. В силу этого при достаточно больших j в общем решении уравнения (13) могут присутствовать сколь угодно большие по абсолютной величине гармоники.

Подобными примерами не исчерпываются недостатки непосредственной формальной аппроксимации. В более сложных ситуациях для порождаемых ею дискретных моделей могут не выполняться даже фундаментальные свойства, присущие исходным объектам.

Проиллюстрируем это, рассмотрев следующую задачу для стационарного уравнения теплопроводности:

$$\frac{d}{dx}\left(k(x)\frac{du}{dx}\right)=0,\quad 0<x<1,\quad k(x)\geqslant C_1>0, \tag{15}$$
$$u(0)=1,\quad u(1)=0.$$

Модель (15) — одна из простейших в теории теплопередачи за тем исключением, что коэффициент теплопроводности $k(x)$ может быть разрывной функцией (теплопроводящий материал состоит из разных веществ).

Раскроем дифференциальный оператор (15):

$$\frac{d}{dx}\left(k(x)\frac{du}{dx}\right)=k\frac{d^2k}{dx^2}+\frac{dk}{dx}\cdot\frac{du}{dx},$$

и используем для полученного выражения совершенно естественные, на первый взгляд, замены (сетка равномерная)

$$\frac{d^2u}{dx^2}\approx u_{\bar{x}x},\quad \frac{dk}{dx}\approx\frac{k_{i+1}-k_{i-1}}{2h},\quad \frac{du}{dx}\approx\frac{u_{i+1}-u_{i-1}}{2h}.$$

С учетом граничных условий (15) придем к разностной схеме

$$k_i\frac{y_{i+1}-2y_i+y_{i-1}}{h^2}+\frac{k_{i+1}-k_i}{2h}\cdot\frac{y_{i+1}-y_{i-1}}{2h}=0, \tag{16}$$
$$0<i<N,\quad y_0=1,\quad y_N=0,$$

имеющей порядок аппроксимации $O(h^2)$.

Пусть $k(x)$ — разрывная кусочно-постоянная функция:

$$k(x)=\begin{cases}k_1, & 0<x<\xi,\\ k_2, & \xi<x<1,\end{cases} \tag{17}$$

где ξ — иррациональное число, $\xi=x_n+\theta h$, $x_n=nh$, $0<\theta<1$, $k_1\neq k_2$.

При такой функции $k(x)$ решение (15), очевидно, линейным образом зависит от x, причем наклон функции $u(x)$ различен в областях с разными значениями коэффициента теплопроводности. Единственное решение определяется из условий сопряжения в точке $x=\xi$, т. е. из условия непрерывности температуры $u^-(\xi)=u^+(\xi)$ и потока тепла $W^-(\xi)=W^+(\xi)$ $(W(x)=-k(x)\,du/dx)$:

$$u(x)=\begin{cases}1-\alpha_0x, & 0\leqslant x\leqslant\xi, \quad \alpha_0=(\varkappa+(1-\varkappa)\xi)^{-1},\\ \beta_0(1-x), & \xi\leqslant x\leqslant 1, \quad \beta_0=\varkappa\alpha_0,\quad \varkappa=k_1/k_2.\end{cases} \tag{18}$$

Нетрудно найти решение разностной задачи (16) в случае (17), пользуясь тем, что уравнение (16) имеет вид $y_{i-1} - 2y_i + y_{i+1} = 0$ при $i \neq n$, $i \neq n+1$ и, следовательно,

$$y_i = y(x_i) = \begin{cases} 1 - \alpha x_i, & 0 \leqslant x \leqslant x_n, \\ \beta(1 - x_i), & x_{n+1} \leqslant x \leqslant 1. \end{cases} \tag{19}$$

Коэффициенты α, β получаются из уравнений (16) в точках $i = n$, $i = n+1$ $(x_n = \xi - \theta h,\ x_{n+1} = \xi + (1-\theta)h)$:

$$\alpha = \frac{1}{\mu + (1-\mu)\xi + h(\lambda - \theta - (1-\theta)\mu)},$$

$$\mu = \frac{3+\varkappa}{5-\varkappa}, \quad \lambda = \frac{5\varkappa - 1}{3\varkappa + 1}, \quad \beta = \mu\alpha.$$

Из этих формул видно, что $\alpha \to \bar{\alpha}_0 = (\mu + (1-\mu)\xi)^{-1}$, $\beta \to \bar{\beta}_0 = \mu\alpha_0$ при $h \to 0$. Поэтому при $h \to 0$ функция $\bar{y}(x,h)$ — решение (9), доопределенное между узлами сетки с помощью линейной интерполяции, — имеет предел

$$\bar{u}(x) = \lim_{h\to 0} \bar{y}(x,h) = \begin{cases} 1 - \bar{\alpha}_0 x, & 0 \leqslant x \leqslant \xi, \\ \bar{\beta}_0(1 - x), & \xi \leqslant x \leqslant 1. \end{cases} \tag{20}$$

Функции (20) и (19) совпадают лишь в случае $\varkappa = 1$ $(k_1 = k_2)$. Следовательно, решение разностной задачи стремится при $h \to 0$ не к решению задачи (13), а к совсем другой функции: разностная схема (16) *расходится*.

Причина ее расходимости устанавливается из анализа решения (20) в точке $x = \xi$: температура в ней непрерывна $(u^-(\xi) = u^+(\xi))$, а поток тепла претерпевает разрыв $(W^-(\xi) \neq W^+(\xi)$; упр. 5). Схема (16) нарушает закон сохранения (баланс) тепла в веществе, т. е. фундаментальный закон, на основе которого получаются все модели процесса теплопередачи. Подобные дискретные модели называются *неконсервативными* и, как правило, не могут использоваться для исследования исходных моделей.

3. Интегро-интерполяционный метод. Из приведенных примеров ясно, что переход к дискретным моделям нельзя осуществлять чисто формально. Он должен оставлять в силе по возможности наибольшее число основных свойств, которыми обладают модели исходных объектов.

Применительно к уравнению (15) это означает, что для отвечающей ему разностной схемы необходимо выполнение дискретного аналога закона сохранения энергии. Широко используемый подход к построению таких схем основан на записи этого закона в интегральной форме для ячеек $x_{i-1} \leqslant x < x_i$, $x_i = ih$, $i = 1, \ldots, N$, разностной сетки с последующей заменой получающихся интегралов и производных приближенными разностными выражениями (*интегро-интерполяционный метод*).

Для стационарного процесса теплопроводности без поглощения и выделения энергии уравнение баланса тепла на отрезке $x_{i-1/2} < x < x_{i+1/2}$ (выбраны полуцелые индексы) означает равенство потоков на его границах:

$$W_{i-1/2} - W_{i+1/2} = 0, \qquad i = 1, \ldots, N-1. \tag{21}$$

Проинтегрируем равенство $W(x) = -k(x)\,du/dx$ на отрезке $x_{i-1} \leqslant x \leqslant x_i$:

$$u_{i-1} - u_i = \int_{x_{i-1}}^{x_i} \frac{W(x)}{k(x)}\,dx,$$

и, предполагая $W(x) = \bar{W}_{i-1/2} = \text{const}$ при $x_{i-1} \leqslant x \leqslant x_i$ (простейшая интерполяция), получим

$$u_{i-1} - u_i \approx \bar{W}_{i-1/2} \int_{x_{i-1}}^{x_i} \frac{dx}{k(x)},$$

откуда приближенное значение для $\bar{W}_{i-1/2}$ дается формулой

$$\bar{W}_{i-1/2} = a_i \frac{u_i - u_{i-1}}{h}, \quad a_i = \left(\frac{1}{h}\int_{x_{i-1}}^{x_i} \frac{dx}{k(x)}\right)^{-1}.$$

Подставив ее в (21), придем к *консервативной* (ср. с (16)) *разностной схеме*

$$\frac{1}{h}\left(a_{i+1}\frac{y_{i+1} - y_i}{h} - a_i \frac{y_i - y_{i-1}}{h}\right) = 0, \qquad 1 \leqslant i \leqslant N-1,$$

$$a_i = \left[\frac{1}{h}\int_{x_{i-1}}^{x_i} \frac{dx}{k(x)}\right]^{-1} = \left[\int_{-1}^{0} \frac{ds}{k(x_i + sh)}\right]^{-1}, \tag{22}$$

для которой закон сохранения энергии выполнен как для каждой ячейки, так и для всего отрезка $[0,1]$ (упр. 6). Таким же образом строятся консервативные дискретные модели для более сложных, в том числе нелинейных и нестационарных процессов теплопередачи.

Применимость интегро-интерполяционного и подобных ему методов распространяется на широкие классы моделей. Он важен, в частности, для задач газовой динамики, решения которых (см. п. 2 § 1) могут быть разрывными. Построим консервативные разностные аппроксимации для уравнений одномерного течения газа, записанных в массовых координатах (п. 5 § 4 гл. II) в дивергентном виде:

$$\frac{\partial \eta}{\partial t} = \frac{\partial v}{\partial m}, \quad \frac{\partial v}{\partial t} = -\frac{\partial p}{\partial m}, \quad \frac{\partial}{\partial t}\left(\varepsilon + \frac{v^2}{2}\right) = -\frac{\partial}{\partial m}(pv), \tag{23}$$

$$0 < t \leqslant T, \quad 0 < m < M_0.$$

Здесь t — время, m — массовая координата; $\eta = 1/\rho$ — удельный объем (ρ — плотность), v — скорость, p — давление и $\varepsilon = \varepsilon(\eta, p)$ — внутренняя энергия газа.

Введем сетку с равномерными шагами по времени и пространству:

$$\bar\omega_h = \{m_i = ih,\ i = 0, 1, \ldots, N;\ h = M_0/N\},$$

$$\omega_\tau = \{t^j = j\tau,\ j = 0, 1, \ldots, M;\ \tau = T/M\}, \quad \bar\omega_{h\tau} = \bar\omega_h \cup \omega_\tau.$$

Ради простоты для сеточных аналогов газодинамических величин сохраним те же обозначения, отнеся функцию v к целым точкам сетки ($m = m_i$), а p, η, ε — к полуцелым ($m = m_{i+1/2}$).

Проинтегрируем второе уравнение (23) в прямоугольнике $m_{i-1/2} \leqslant m \leqslant m_{i+1/2}$, $t^j \leqslant t \leqslant t^{j+1}$:

$$\int\limits_{m_{i-1/2}}^{m_{i+1/2}} (v^{j+1} - v^j)\, dm + \int\limits_{t^j}^{t^{j+1}} (p_{i+1/2} - p_{i-1/2})\, dt = 0,$$

а остальные — в прямоугольнике $m_i \leqslant m \leqslant m_{i+1}$, $t^j \leqslant t \leqslant t^{j+1}$:

$$\int\limits_{m_i}^{m_{i+1}} (\eta^{j+1} - \eta^j)\, dm - \int\limits_{t^j}^{t^{j+1}} (v_{i+1} - v_i)\, dt = 0,$$

$$\int\limits_{m_i}^{m_{i+1}} \left[\left(\varepsilon + \frac{v^2}{2}\right)^{j+1} - \left(\varepsilon + \frac{v^2}{2}\right)^{j}\right] dm + \int\limits_{t^j}^{t^{j+1}} [(pv)_{i+1} - (pv)_i]\, dt = 0.$$

Временные интегралы, входящие в эти тождества, заменим выражениями

$$\int\limits_{t^j}^{t^{j+1}} p\, dt \approx p^{(\sigma_1)}\tau, \quad \int\limits_{t^j}^{t^{j+1}} v\, dt \approx v^{(\sigma_2)}\tau, \quad \int\limits_{t^j}^{t^{j+1}} (pv)_i\, dt \approx p_{*i}^{(\sigma_3)} v_i^{(\sigma_4)}\tau,$$

где $f^{(\sigma_\alpha)} = \sigma_\alpha f^{j+1} + (1 - \sigma_\alpha) f^j$, σ_α — веса ($\alpha = 1, 2, 3, 4$), $p_{*i} = 0{,}5\,(p_{i-1/2} + p_{i+1/2})$. Пространственные интегралы заменим по очевидному правилу, например,

$$\int\limits_{m_{i-1/2}}^{m_{i+1/2}} v\, dm \approx v_i h, \quad \int\limits_{m_i}^{m_{i+1}} \eta\, dm \approx \eta_{i+1/2} h.$$

В итоге придем к четырехпараметрическому семейству разностных схем

$$\begin{aligned}
&\frac{v_i^{j+1} - v_i^j}{\tau} = -\left(\frac{p_{i+1/2} - p_{i-1/2}}{h}\right)^{(\sigma_1)},\\
&\frac{\eta_{i+1/2}^{j+1} - \eta_{i+1/2}^{j}}{\tau} = \left(\frac{v_{i+1} - v_i}{h}\right)^{(\sigma_2)},\\
&\frac{E_{i+1/2}^{j+1} - E_{i+1/2}^{j}}{\tau} = \frac{p_{*i+1}^{(\sigma_3)} v_{i+1}^{(\sigma_2)} - p_{*i}^{(\sigma_3)} v_i^{(\sigma_4)}}{h},
\end{aligned} \tag{24}$$

где $E_{i+1/2} = \varepsilon_{i+1/2} + (v_i^2 + v_{i+1}^2)/2$ — полная энергия i-й ячейки газа. Они представляют собой консервативные дискретные аналоги уравнений (23) при любых σ_α, $\alpha = 1, \ldots, 4$. Для них имеют место разностные аналоги законов сохранения массы, импульса и полной энергии в любой ячейке сетки (интегрированием уравнений (24) по сетке $\bar\omega_{h\tau}$ нетрудно установить справедливость законов сохранения в дискретной форме для всей массы газа $0 \leqslant m \leqslant M_0$).

При $\sigma_1 = 0$, $\sigma_2 = \sigma_3 = \sigma_4 = 1$ система (24) может быть решена по явным формулам: сначала находятся v_i^{j+1}, затем $\eta_{i+1/2}^{j+1}$, а из третьего уравнения (24) и уравнения состояния $p\eta = (\gamma - 1)\varepsilon$ (в случае идеального газа) методом прогонки определяются с использованием краевых условий при $i = 0$, $i = N$ значения $p_{i+1/2}$ для всех $i = 0, 1, \ldots, N$. Неявные схемы (24) решаются с помощью итерационных методов.

В отличие от уравнений параболического типа, явные схемы для гиперболических уравнений устойчивы не при условии $\tau \leqslant Ch^2$, а при выполнении гораздо более мягкого неравенства $\tau \leqslant Ch$, и поэтому часто применяются на практике. Заметим также, что при практических вычислениях в газодинамические разностные схемы вводится *искусственная «вязкость»*, сглаживающая сильные разрывы. Это заметно облегчает проведение расчетов, так как отпадает необходимость в специальном выделении областей разрывов (*однородные схемы*).

4. Принцип полной консервативности. Более существенное отличие дискретных моделей газовой динамики от моделей теплопередачи заключается в разнообразии способов математического представления законов сохранения для газа. В частности, третье уравнение (23) может быть записано не для полной энергии $E = \varepsilon + v^2/2$, а для внутренней энергии ε, причем по меньшей мере в двух видах:

$$\frac{\partial \varepsilon}{\partial t} = -p\frac{\partial v}{\partial m}, \quad \frac{\partial \varepsilon}{\partial t} = -p\frac{\partial \eta}{\partial t}. \tag{25}$$

Для исходной модели эти и другие формы записи эквивалентны и переходят одна в другую при соответствующих преобразованиях уравнений. В случае дискретных моделей данное свойство отнюдь не гарантируется. Например, из третьего уравнения консервативной схемы (24), вообще говоря, не следуют дискретные аналоги уравнений

(25). Верно и обратное: для схемы, аппроксимирующей уравнения внутренней энергии, не обязательно выполняется закон сохранения полной энергии. Простым примером служит неконсервативная схема «крест»

$$\frac{v_i^{j+1}-v_i^j}{\tau}=-\frac{p_{i+1/2}^{j+1/2}-p_{i-1/2}^{j+1/2}}{h},$$

$$\frac{\eta_{i+1/2}^{j+3/2}-\eta_{i+1/2}^{j+1/2}}{\tau}=\frac{v_{i+1}^{j+1}-v_i^{j+1}}{h}, \qquad (26)$$

$$\frac{\varepsilon_{i+1/2}^{j+3/2}-\varepsilon_{i+1/2}^{j+1/2}}{\tau}=-p_{i+1/2}^{j+3/2}\frac{v_{i+1}^{j+1}-v_i^j}{h},$$

для написания которой используется «шахматная» сетка (см. рис. 82; точками и крестиками обозначены узлы применяемого шеститочечного шаблона; функции ε, p, η отнесены к полуцелым, а v — к целым узлам) и легко решаемая в случае $p\eta=(\gamma-1)\varepsilon$ по явным формулам.

Подобная «частичная», неполная консервативность дискретных моделей может сделать их непригодными для численного моделирования (сохранение полной энергии не означает, что ее составляющие — кинетическая и внутренняя энергии, а с ними скорость и температура газа, — рассчитываются правильно, и т. п.). Поэтому при построении дискретных моделей необходимо отразить в них как можно большее число фундаментальных свойств исходных объектов.

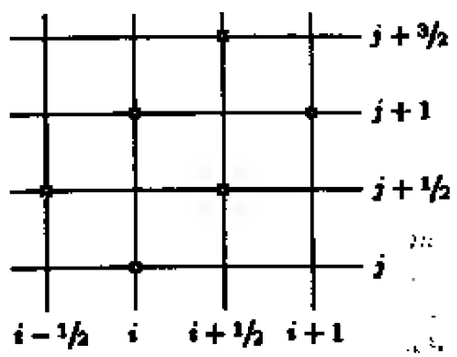


Рис. 82

Взяв за исходное четырехпараметрическое семейство консервативных схем (24), получим обладающие сформулированным качеством разностные аппроксимации уравнений (23).

Для упрощения дальнейших выкладок введем обозначения

$$\bar p_i=p_{i+1/2},\quad \bar\eta_i=\eta_{i+1/2},\quad \bar\varepsilon_i=\varepsilon_{i+1/2},\quad \bar p=\bar p_i^j,\quad v=v_i^j,$$

$$\frac{p_{i+1/2}-p_{i-1/2}}{h}=\bar p_m,\quad \frac{v_{i+1}-v_i}{h}=v_m,$$

после чего черту над p, η и ε будем опускать (кроме специальных случаев). С той же целью будем вместо третьего уравнения (24) рассматривать аппроксимацию первого уравнения (25) для внутренней энергии. Кроме того, понадобится формула

$$f^{(\beta)}=f^{(\alpha)}+\tau(\beta-\alpha)f_t, \qquad (27)$$

где α и β — любые числа, $f^{(\alpha)}=\alpha\hat f+(1-\alpha)f$, $\hat f=f^{j+1}$.

С учетом сделанных замечаний из (24) получается четырехпараметрическое семейство схем

$$v_t=-p_m^{(\sigma_1)},\quad \eta_t=v_m^{(\sigma_2)},\quad \varepsilon_t=-p^{(\sigma_3)}v_m^{(\sigma_4)}, \qquad (28)$$

от которых потребуем *полной консервативности*, т. е. того, чтобы они аппроксимировали также уравнение для полной энергии (третье уравнение (23)) и второе уравнение (25).

Из второго уравнения (28), равенства $v_m^{(\sigma_4)}=v_m^{(\sigma_2)}-\tau(\sigma_4-\sigma_2)v_{mt}=\eta_t+\tau(\sigma_4-\sigma_2)v_{mt}$ и третьего уравнения (28) следует

$$\varepsilon_t=-p^{(\sigma_3)}\eta_t+\delta_1E,$$

где $\delta_1E=-\tau(\sigma_4-\sigma_2)p_{mt}^{(\sigma_3)}$. Таким образом, схема (28) аппроксимирует второе уравнение (25) лишь при выполнении равенства $\sigma_2=\sigma_4$. Если это равенство не выполняется, то возникает *дисбаланс* δ_1E внутренней энергии, вызванный появлением в дискретной среде дополнительных (фиктивных) источников и стоков энергии.

Найдем теперь δ_2E — дисбаланс полной энергии. Умножим первое уравнение (28) на $v^{(0,5)}=0{,}5(v+\hat v)$ и получим уравнение

$$\frac{v_t^2}{2}=-(v^{0,5})\cdot p_m^{(\sigma_1)},$$

после чего сложим его с третьим уравнением (28):

$$\left(\varepsilon+\frac{v^2}{2}\right)_t=-p^{(\sigma_3)}v_m^{(\sigma_4)}-v^{0,5}p_m^{(\sigma_1)}. \qquad (29)$$

Преобразуем правую часть (29) с помощью формулы (27):

$$p^{(\sigma_3)}v_m^{(\sigma_4)}+v^{0,5}p_m^{(\sigma_1)}=$$
$$=(p^{(\sigma_1)}+\tau(\sigma_3-\sigma_1)p_t)(v_m^{(0,5)}+\tau(\sigma_4-0{,}5)v_{mt})+v^{(0,5)}p_m^{(\sigma_1)}=$$
$$=(p_{(-1)}^{(\sigma_1)}v^{0,5})_m+\delta_2E,$$

где приняты обозначения

$$p_{(-1)}=\bar p_{i-1}=p_{i-1/2},$$

$$\delta_2E=\tau(\sigma_3-\sigma_1)v_m^{(0,5)}p_t+$$
$$+\tau(\sigma_4-0{,}5)p_{mt}^{(\sigma_1)}+\tau^2(\sigma_3-\sigma_1)(\sigma_4-0{,}5)p_tv_{mt}.$$

В итоге (29) принимает вид

$$\left(\varepsilon+\frac{v^2}{2}\right)_t=-(p_{(-1)}^{(\sigma_1)}v^{(0,5)})_m-\delta_2E.$$

Таким образом, уравнение (29) не аппроксимирует уравнение для полной энергии (то же самое относится и к схеме (26); см. упр. 7). Дисбаланс δ_2E, как и дисбаланс δ_1E, имеет искусственное происхождение.

Он отсутствует лишь в случае $\sigma_3 = \sigma_1$, $\sigma_4 = 0{,}5$ и, следовательно, $\sigma_2 = 0{,}5$.

Объединяя эти результаты, придем вместо четырехпараметрического семейства схем (28), полученных с помощью интегро-интерполяционного метода, к однопараметрическому семейству

$$v_t = -p_m^{(\sigma_1)}, \quad \eta_t = v_m^{(0,5)}, \quad \varepsilon_t = -p^{(\sigma_1)} v_m^{(0,5)} \tag{30}$$

полностью консервативных ($\delta_1 E = \delta_2 E = 0$) дискретных моделей газовой динамики.

Заметим, что третье из уравнений (30) можно заменить одним из уравнений

$$\varepsilon_t = -p^{(\sigma_1)}\eta_t, \quad \left(\varepsilon + \frac{v^2}{2}\right)_t = -\left(p_{(-1)}^{(\sigma_1)} v^{(0,5)}\right)_m,$$

а вместо последнего взять одно из уравнений

$$\left(\varepsilon + \frac{v_{(+1)}^2}{2}\right)_t = -\left(p^{(\sigma_1)} v^{(0,5)}\right)_m, \quad \left(\varepsilon + \frac{v^2 + v_{(+1)}^2}{4}\right)_t = -\left(p_{\circ}^{(\sigma_1)} v^{(0,5)}\right)_m.$$

Схема (30) в общем случае обладает аппроксимацией $O(\tau + h^2)$, а при $\sigma_1 = 0{,}5$ аппроксимация по времени имеет второй порядок. При $\sigma_1 = 0$ схема (30) явная, и ее решение находится по простым формулам.

Из структуры $\delta_1 E$ и $\delta_2 E$ хорошо видно, что они особенно велики, если характеристики газа сильно меняются по пространству и по времени (разрывные течения, сильная неоднородность среды, режимы с обострением и т. п.). В этих ситуациях фиктивные источники энергии становятся сравнимыми с реальными величинами, и результаты численного моделирования исходного объекта заметно отличаются от его действительного поведения. Дисбалансы могут быть уменьшены с помощью уменьшения шага по времени τ, что, естественно, приводит к соответствующему увеличению числа расчетных операций (уменьшение шага сетки по пространству вообще не сказывается на величине $\delta_1 E$, $\delta_2 E$).

Принцип полной консервативности, нашедший применение при решении многих сложных задач, — один из надежных подходов к построению дискретных моделей с необходимыми качествами.

5. Построение разностных схем с помощью вариационных принципов. Вариационные принципы, служащие одним из основных способов получения математических моделей различных объектов, широко используются также и для построения соответствующих дискретных аналогов. Этому способствуют их универсальность, относительная простота применения, связь с законами сохранения и свойствами симметрии. Самый естественный путь при таком подходе — дискретизация исходного объекта (среды), формулировка для полученного объекта вариационного принципа и вывод на его основе связей (уравнений) для дискретных величин, т. е. завершение конструирования искомой дискретной модели.

Построим, следуя этой логике, разностные схемы для уравнений (23) одномерной газовой динамики в лагранжевых координатах и сравним их со схемами пп. 3, 4.

Будем рассматривать газ как совокупность соседствующих друг с другом материальных «точек» («частиц»), разбив его массу M_0 для простоты равномерно на N частей с массой $m_{i+1/2} = M_0/N$, $i = 0, 1, \ldots, N-1$ (масса $m_{i+1/2}$ — аналог шага h в (24), (25)). Другими словами, как и в п. 3, введем разностную сетку с узлами в точках $i = 0, 1, \ldots, N$ и равномерными по массе ячейками. Координаты и скорости частицы (ячейки) охарактеризуем через декартовы координаты x_i, x_{i+1} и скорости v_i, v_{i+1} соответствующих узлов. Остальные параметры материальных точек — плотность $\rho_{i+1/2}$, удельный объем $\eta_{i+1/2}$, давление $p_{i+1/2}$, внутреннюю энергию $\varepsilon_{i+1/2}$ — отнесем к середине ячейки (полуцелый индекс).

Применим для построенной таким образом дискретной среды принцип Гамильтона, выбрав $x_i(t)$, $x_{i+1}(t)$ в качестве обобщенных координат ячеек (величины $v_i(t) = dx_i/dt = \dot{x}_i$, $v_{i+1}(t) = dx_{i+1}/dt = \dot{x}_{i+1}$ играют роль обобщенных скоростей). Кинетическую энергию системы определим достаточно очевидным выражением

$$T = \sum_{i=0}^{N-1} T_{i+1/2} = \sum_{i=0}^{N-1} m_{1+1/2} \frac{v_i^2 + v_{i+1}^2}{4}$$

как сумму энергий каждой частицы (возможны и другие выражения)

$$T_{i+1/2} = \frac{m_{i+1/2}}{2} \frac{v_i^2 + v_{i+1}^2}{2}.$$

«Потенциальную» энергию частиц будем соотносить с их внутренней энергией $\varepsilon_{i+1/2} = \varepsilon_{i+1/2}(x_i, x_{i+1})$, поскольку ее «высвобождение» дает возможность совершить работу, подобно тому как замыкание обкладок заряженного конденсатора высвобождает запасенную в нем энергию, преобразующуюся в движение зарядов; см. п. 4 § 2 гл. III. Заметим, что такое определение потенциальной энергии для рассматриваемой системы, как и обоснование потенциальности ее движения, требует подробного и громоздкого анализа (в п. 3 § 2 гл. III он проведен для более простого случая малых колебаний струны). Поэтому ограничимся приведенными выше качественными соображениями.

Суммарная потенциальная энергия совокупности ячеек есть

$$V = \sum_{i=0}^{N-1} V_{i+1/2} = \sum_{i=0}^{N-1} m_{i+1/2}\varepsilon_{i+1/2},$$

где $V_{i+1/2}$ — энергия отдельной частицы.

«Лагранжиан» дискретной газовой среды дается выражением

$$L = T - V = \sum_{i=0}^{N-1} m_{i+1/2}\left(\frac{v_i^2 + v_{i+1}^2}{4} - \varepsilon_{i+1/2}\right), \tag{31}$$

а функционал действия — формулой

$$Q = \int_{t_0}^{t_1} L(x(t), \dot{x}(t), t)\, dt =$$

$$= \int_{t_0}^{t_1} \sum_{i=0}^{N-1} m_{i+1/2} \left(\frac{v_i^2 + v_{i+1}^2}{4} - \varepsilon_{i+1/2}(x_i, x_{i+1}) \right) dt. \qquad (32)$$

Проварьируем в соответствии со схемой применения принципа Гамильтона (п. 3 § 1 гл. III) действие (32) по всем кинематически возможным путям, вычислив сначала вариации составляющих его членов. Так как $v_i = \dot{x}_i$, то

$$\delta\left(\frac{v_i^2}{4}\right) = \frac{1}{4}\frac{\partial v_i^2}{\partial \dot{x}_i}\,\delta\dot{x}_i = \frac{1}{2}\, v_i\, \delta\dot{x}_i.$$

Вариация внутренней энергии дается формулой

$$\delta\varepsilon_{i+1/2} = \frac{\partial \varepsilon_{i+1/2}}{\partial x_i}\,\delta x_i + \frac{\partial \varepsilon_{i+1/2}}{\partial x_{i+1}}\,\delta x_{i+1},$$

а для вариации действия имеем

$$\delta Q = \int_{t_0}^{t_1} m_{i+1/2} \sum_{i=0}^{N-1} \left(\frac{1}{2}\, v_i\, \delta\dot{x}_i + \frac{1}{2}\, v_{i+1}\, \delta\dot{x}_{i+1} - \right.$$

$$\left. - \frac{\partial \varepsilon_{i+1/2}}{\partial x_i}\,\delta x_i - \frac{\partial \varepsilon_{i+1/2}}{\partial x_{i+1}}\,\delta x_{i+1} \right) dt.$$

Проинтегрируем в последнем равенстве члены $v_i\,\delta\dot{x}_i$ и $v_{i+1}\,\delta\dot{x}_{i+1}$ по частям, учтем перестановочность операций варьирования и дифференцирования по времени ($\delta\dot{x}_i = d(\delta x_i)/dt$) и примем во внимание условие $\delta x_i = 0$, $i = 0, \ldots, N-1$, при $t = t_0$, $t = t_1$. Получим выражение, из которого видно, что в силу независимости вариаций δx_i для δQ имеем, что $\delta Q = 0$ лишь тогда, когда равны нулю коэффициенты при любой δx_i, т. е. при выполнении уравнений

$$\frac{dv_i}{dt} = -\frac{\partial \varepsilon_{i+1/2}}{\partial x_i} - \frac{\partial \varepsilon_{i-1/2}}{\partial x_i}, \qquad i = 0, \ldots, N-1.$$

В их левых частях записано ускорение ячеек газа. Чтобы выяснить смысл правых частей, напомним, что для рассматриваемых здесь адиабатических течений справедливы равенства $d\varepsilon = -p\,d\eta$ и $\varepsilon(p, \eta) = \varepsilon(\eta)$ (см. формулы (27), (29) из § 4 гл. II). Отсюда для дискретной среды получаем $d\varepsilon_{i+1/2} = -p_{i+1/2}\,d\eta_{i+1/2}$, $\varepsilon_{i+1/2} = \varepsilon(\eta_{i+1/2}) = \varepsilon[(x_{i+1} - x_i)/m_{i+1/2}]$. Проводя преобразования

$$\frac{\partial \varepsilon_{i+1/2}}{\partial x_i} = \frac{\partial \varepsilon_{i+1/2}}{\partial \eta_{i+1/2}} \cdot \frac{\partial \eta_{i+1/2}}{\partial x_i} = \frac{p_{i+1/2}}{m_{i+1/2}},$$

$$\frac{\partial \varepsilon_{i-1/2}}{\partial x_i} = \frac{\partial \varepsilon_{i-1/2}}{\partial \eta_{i-1/2}} \cdot \frac{\partial \eta_{i-1/2}}{\partial x_i} = -\frac{p_{i-1/2}}{m_{i-1/2}},$$

придем к окончательному виду полученных из принципа Гамильтона уравнений движения газовых «частиц» (на равномерной сетке):

$$\frac{dv_i}{dt} = -\frac{p_{i+1/2} - p_{i-1/2}}{(m_{i+1/2} + m_{i-1/2})/2}, \qquad i = 0, \ldots, N-1. \qquad (33)$$

Присоединим к (33) уравнение

$$\frac{d\eta_{i+1/2}}{dt} = \frac{v_{i+1} - v_i}{m_{i+1/2}}, \qquad i = 0, \ldots, N-1, \qquad (34)$$

следующее из равенства $\eta_{i+1/2} = (x_{i+1} - x_i)/m_{i+1/2}$ (закон сохранения массы ячейки), а также вытекающее из равенства $d\varepsilon_{i+1/2} = -p_{i+1/2}\,d\eta_{i+1/2}$ и из (34) уравнение

$$\frac{d\varepsilon_{i+1/2}}{dt} = -p_{i+1/2}\frac{d\eta_{i+1/2}}{dt} = -p_{i+1/2}\frac{v_{i+1} - v_i}{m_{i+1/2}}, \qquad i = 0, \ldots, N-1, \quad (35)$$

выражающее закон изменения внутренней энергии.

Вместе с заданным уравнением состояния $\varepsilon_{i+1/2} = \varepsilon(\eta_{i+1/2})$ система (33)–(35) представляет собой *полудискретную* модель динамики газа, отвечающую всем необходимым законам сохранения — импульса, массы, энергии (заметим, что закон сохранения полной энергии $E = T + V$ сразу следует из инвариантности лагранжиана (31) по отношению к сдвигу времени).

Проводя в (33)–(35) замену производных по времени конечными разностями, переходя к безындексным обозначениям и вводя веса σ_1, σ_2, σ_3, σ_4, получим дискретную модель газовой динамики

$$v_t = -p^{(\sigma_1)}\bar{m}, \quad \eta_t = v_m^{(\sigma_2)}, \quad \varepsilon_t = -p^{(\sigma_3)} v_m^{(\sigma_4)},$$

совпадающую с разностной схемой (28), построенной в п. 3 с помощью интегро-интерполяционного метода (при соответствующем выборе весов она становится полностью консервативной).

Применимость продемонстрированного подхода не ограничивается рассмотренным выше простейшим случаем (см., например, упр. 8). Вариационные принципы эффективно используются для получения дискретных моделей в весьма трудных ситуациях (многомерные процессы, сетки сложной структуры и т. д.).

6. Использование иерархического подхода к получению дискретных моделей. Основная идея данного подхода — использование знания о месте подлежащей дискретизации модели в иерархии моделей изучаемого объекта. Если иерархия построена по принципу «сверху вниз», то, проведя частичную или полную дискретизацию общей модели, затем переходят к дискретной модели более низкого уровня. Естественно, используется способ перехода, принятый для исходных моделей. В результате выполнения подобной процедуры в построенную дискретную модель привносятся дополнительные черты модели

более высокого уровня, и это может повысить адекватность аппроксимирующей модели исходному объекту.

Конкретизируем эти рассуждения, рассмотрев уравнения, описывающие одномерные течений невязкого нетеплопроводного газа (записанную в дивергентном виде одномерную систему (4), (10), (15) из § 4 гл. II)

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho V)}{\partial x} = 0,$$
$$\frac{\partial(\rho V)}{\partial t} + \frac{\partial}{\partial x}(\rho V^2 + p) = 0, \qquad (36)$$
$$\frac{\partial E}{\partial t} + \frac{\partial}{\partial x}[V(E+p)] = 0.$$

Здесь $E = \rho(V^2/2 + 3RT/2)$ — сумма кинетической и внутренней энергий единицы объема газа, т. е. его полная энергия, $T = p/(\rho R)$ — температура, $R = k/m$ — газовая постоянная, k — постоянная Больцмана, m — масса атома или молекулы газа.

Напомним, что в пп. 4, 5 § 3 гл. III уравнения (36) получены из кинетического уравнения Больцмана в результате его «осреднения» при условии, что функция распределения $f(x, v, t)$ локально максвелловская (формула (5) § 3 гл. III):

$$f^{(0)}(v) = n\left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left[-\frac{m}{2kT}(v - V)^2\right], \qquad (37)$$

где v — скорость хаотического движения частиц газа (при максвелловском распределении в среде отсутствуют вязкие напряжения и потоки тепла).

Дискретные аналоги уравнений (36) можно строить любым из описанных в предыдущих пунктах способом. Иерархический подход в данном случае (и для некоторых обобщений (36)) заключается в следующем:

1) функция распределения f (решение уравнения (2) § 3 гл. III) и макрохарактеристики газа рассматриваются как функции дискретного аргумента t^j, $j = 0, 1, \ldots$, т. е. проводится дискретизация по времени исходной модели — кинетического уравнения Больцмана;

2) величина $\tau = t^{j+1} - t^j$ (шаг дискретизации) берется достаточно малой, чтобы удовлетворить условию $\tau \leqslant l/|\langle v\rangle|$, где l — характерная длина свободного пробега, $|\langle v\rangle|$ — модуль характерной скорости частиц. Поэтому за время τ число столкновений между частицами невелико и ими можно пренебречь (см. вывод уравнения Больцмана);

3) функция распределения $f^{j+1} = f(x, v, t^{j+1})$ изменяется (см. вывод уравнения Больцмана) в сравнении с функцией $f^j = f(x, v, t^j)$ лишь за счет изменения частицами их фазового объема, поэтому справедлива простая связь

$$f(x, v, t^{j+1}) = f(x - v\tau, v, t^j), \qquad (38)$$

являющаяся решением уравнения Больцмана в отсутствие столкновений;

4) связь (38) упрощается путем разложения f^{j+1} в ряд Тейлора по параметру $v\tau$:

$$f^{j+1} = f^j - \tau\frac{\partial f^j}{\partial x}v + \frac{\tau^2}{2}\frac{\partial^2 f^j}{\partial x^2}v^2 + \ldots \qquad (39)$$

с точностью до членов третьего порядка; так как функция распределения экспоненциально убывает с ростом v^2, то вклад частиц с большими скоростями считается малым и соответствующие члены в (39) отбрасываются;

5) подобно тому, как это делалось в п. 4 § 3 гл. III при выводе уравнений для моментов, соотношения (39) последовательно умножаются на функции $\Phi(v)$, соответственно равные 1, mv, $mv^2/2$ (*сумматорные инварианты*), и интегрируются по всем скоростям v; в результате связь между f^{j+1} и f^j преобразуется в связи уже между средними гидродинамическими параметрами газа такими, как плотность, поток массы, полная энергия:

$$\rho = \int mf\,dv, \quad \rho V = \int mvf\,dv, \quad E = \int \frac{mv^2}{2}f\,dv,$$

и другими величинами, взятыми в моменты t^j, t^{j+1} (иными словами, получается *дискретная по времени* гидродинамическая модель среды);

6) после соответствующей дискретизации по пространству строятся окончательные дискретные (и по времени, и по пространству) модели газа, учитывающие некоторые черты исходного кинетического уравнения (*кинетически согласованные разностные схемы*); они, вообще говоря, не могут быть получены аппроксимацией гидродинамических моделей (например, уравнений (36)), поскольку для их конструирования используются свойства модели более высокого иерархического уровня.

Наиболее простой вариант описанных выше процедур соответствует случаю локально максвелловской функции распределения (37) и отвечающих ей (для одномерных течений) уравнений (36), полудискретные аналоги которых выглядят так:

$$\frac{\hat\rho - \rho}{\tau} + \frac{\partial}{\partial x}(\rho V) = \frac{\tau}{2}\frac{\partial^2}{\partial x^2}(\rho V^2 + p),$$
$$\frac{(\widehat{\rho V}) - \rho V}{\tau} + \frac{\partial}{\partial x}(\rho V^2 + p) = \frac{\tau}{2}\frac{\partial^2}{\partial x^2}(\rho V^3 + 3pV), \qquad (40)$$
$$\frac{\hat E - E}{\tau} + \frac{\partial}{\partial x}[V(E+p)] = \frac{\tau}{2}\frac{\partial^2}{\partial x^2}\left[V^2(E+2p) + \frac{p}{\rho}(E+p)\right],$$

где для упрощения записи введено обозначение $g^{j+1} = \hat g$, $g^j = g$. При их выводе интегралы, получающиеся в правой части (39), представлены (с учетом свойств функции распределения $f^{(0)}$) как суммы и

произведения моментов функции e^{-z^2} различных порядков. Нечетные моменты равны нулю, а четные определяются по формуле

$$\int_{-\infty}^{\infty} e^{-z^2} z^{2r}\, dz = \frac{\pi^{1/2}}{2^r}(2r-1)!!.$$

После проведения тем или иным способом дискретизации пространственных операторов из (40) получаются полностью дискретные аналоги уравнений (36), численное решение которых может быть найдено по явным формулам (нетрудно построить и неявные варианты модели (40), поменяв в (38) местами f^{j+1} и f^j). Стоящие в правой части (40) члены первого порядка малости по τ представляют собой добавки к обычным дискретным моделям для (36), несущие дополнительную информацию о происхождении аппроксимируемой модели.

Не приводя достаточно громоздкий вывод всей системы (40), ограничимся лишь уравнением неразрывности. Представляя (39) в виде

$$m\frac{f^{j+1}-f^j}{\tau} + m\frac{\partial f^j}{\partial x}v = m\frac{\tau}{2}\frac{\partial^2 f^j}{\partial x^2}v^2,$$

умножая это выражение на $\Phi(v)=1$ и интегрируя его по v, получим

$$\frac{\rho^{j+1}-\rho^j}{\tau} + \int m\frac{\partial f^j}{\partial x}v\,dv = \frac{\tau}{2}\int m\frac{\partial^2 f^j}{\partial x^2}v^2\,dv,$$

или, вынося знак дифференцирования за интеграл (см. формулы (10) § 3 гл. III),

$$\frac{\rho^{j+1}-\rho^j}{\tau} + \frac{\partial}{\partial x}\int f^j v\,dv = \frac{\tau}{2}\frac{\partial^2}{\partial x^2}\int f^j v^2\,dv.$$

Преобразуем интегралы в последнем равенстве в соответствии с определениями (см. п. 4 § 3 гл. III) функций V и p и приходим к уравнению

$$\frac{\hat\rho-\rho}{\tau} + \frac{\partial}{\partial x}(\rho V) = \frac{\tau}{2}\frac{\partial^2}{\partial x^2}(\rho V^2 + p),$$

т. е. к первому из уравнений (36).

Описанный выше подход обобщается на многомерные аналоги уравнений (36) и на более сложные модели течений газа, например на уравнения Навье—Стокса (см. п. 5 § 3 гл. III). В последнем случае вместо максвелловской функции распределения $f = f^{(0)}$ используется, естественно, следующее приближение — функция $f = f^{(0)} + f^{(1)}$, отвечающая среде с ненулевыми вязкими напряжениями и потоками тепла (см. п. 4 § 3 гл. III).

Поясним: граничные условия для моделей типа (40) не могут быть взяты в точности такими же, как в случае уравнений Эйлера (или Навье—Стокса) уже хотя бы потому, что при автоматическом их переносе дискретная модель стала бы неконсервативной. Поэтому не только уравнения модели (40), но также и краевые условия для них должны

быть согласованы с краевыми условиями для занимающего более высокий иерархический уровень кинетического уравнения.

Кинетически согласованные разностные схемы в силу своего построения обладают рядом свойств, делающих их весьма эффективными для численного моделирования многих достаточно сложных течений газа.

УПРАЖНЕНИЯ

1. Разлагая функцию $u(x)$ в ряд Тейлора в окрестности точки x и удерживая достаточное число членов, проверьте, что разностная производная $u_{\bar x x}$ приближает производную d^2u/dx^2 со вторым порядком.

2. Пусть в задаче (3) с $y_0 = y_N = 0$ узел $i_0 \neq 0, N$ таков, что $|y_{i_0}| = \max|y_i|$, $0 \leqslant i \leqslant N$. Записывая уравнение (3) в точке i_0, покажите справедливость неравенства $|y_{i_0}| \leqslant |\varphi_{i_0}|/2$ и, тем самым, неравенства (5), где $C = 1/2$.

3. Разлагая функцию $u(x,t)$ в окрестности точки $x = x_i$, $t = (t^{j+1}+t^j)/2$ в ряд Тейлора, убедитесь, что при $\sigma = 1/2$ погрешность аппроксимации схемы (10) равна $O(\tau^2 + h^2)$.

4. Проверьте прямой подстановкой, что задача на собственные значения $X_{\bar x x} + \lambda X = 0$, $0 < x < ih < 1$; $X(0) = X(1)$, $X(x) \not\equiv 0$, имеет нетривиальные решения — собственные функции вида $X^{(k)} = \sqrt{2}\sin \pi k x$.

5. Проверьте, что для решения (20) разность между $W^-(\xi)$ и $W^+(\xi)$ равна $q = -\alpha_0(\mu - \varkappa)k_2$ и что $q \to \pm\infty$ при $\varkappa \to \pm 0$, т. е. мощность фиктивного источника тепла может быть сколь угодно большой.

6. Суммируя уравнение (22) по $i = 1, \ldots, N-1$, получите разностный закон сохранения тепла во всей сеточной области $W_{1/2} - W_{N-1/2} = 0$, где $W_{i-1/2} = -a_i(y_i - y_{i-1})/h$.

7. Введите обозначения $p^{j+1/2}_{i+1/2} = \bar p^j_i = \bar p$, $\eta^{j+3/2}_{i+1/2} = \bar\eta^{j+1}_i = \hat\eta$ и т. д. Запишите схему «крест» (26) в виде

$$v_t = -\bar p_m, \quad \eta_t = \hat v_m, \quad \varepsilon_t = -\hat p \hat v_m.$$

Умножая $v_t = -\bar p_m$ на $v^{(0,5)}$ и повторяя рассуждения, использованные при выводе (30), представьте третье уравнение (26) в виде

$$\left(\varepsilon + \frac{v^2}{2}\right)_t = -\left(\bar p_{-1} v^{(0,5)}\right)_m - \delta E,$$

где дисбаланс полной энергии $\delta E = \tau \bar p_t v_m^{(0,5)} + 0{,}5\,\tau p v_{mt} + 0{,}5\,\tau^2 p_t v_{mt}$, в отличие от случая схемы (24), неустранимы.

8. Проверьте, повторяя в случае неравномерной по массе разностной сетки построения п. 5, что уравнения (34), (35) модели (33)–(35) остаются без изменений, а уравнение (33) хотя и несколько видоизменяется, но сохраняет свой физический смысл (получите его).

Библиография к главе V: [2, 6, 8, 12, 15, 17, 18, 21, 22, 24, 26, 29, 31, 33, 34, 46, 51, 59, 61–65, 67, 68, 70, 71, 74, 75, 80, 87–89].

ГЛАВА VI

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ СЛОЖНЫХ ОБЪЕКТОВ

§ 1. Задачи технологии и экологии

Покажем необходимость широкого применения вычислительного эксперимента для анализа и прогноза крупных технологических и экологических проектов. Приведем конкретные примеры, иллюстрирующие тесную взаимосвязь проблем технологии и экологии, неизбежность использования математического моделирования для их решения.

1. Физически «безопасный» ядерный реактор. Ядерная энергетика служит одной из основ индустрии многих развитых стран, занимая в ряде из них первое, в сравнении с другими видами энергетики, место по выработке электроэнергии. Проблемы ее безопасности сложны и разнообразны: захоронение отходов и уменьшение потребности в добыче урана, эффективное использование обедненного и освобождающегося при ядерном разоружении урана и, конечно же, предотвращение тяжелых аварий на реакторах.

В реакторах любого типа определяющими физическими процессами являются нейтронно-ядерные реакции, приводящие к выделению энергии в его активной зоне, и отвод от этой зоны тепла, используемого затем для получения электроэнергии. Работающий реактор поддерживается в критическом состоянии, когда число выделяющихся нейтронов таково, что вырабатываемая мощность практически не зависит от времени. В подкритическом режиме нейтронов появляется меньше, чем теряется, и реакция деления быстро затухает. В надкритическом состоянии, наоборот, нейтронный выход слишком велик, и это может привести к разогреву и «взрыву» активной зоны.

В обычных реакторах критическое состояние физически неустойчиво, оно поддерживается искусственно с помощью очень сложной системы управления. Без такой системы происходит выход либо на подкритический, либо на надкритический режимы. Реактор делается с запасом реактивности (надкритичность), который компенсируется введением в активную зону специальных стержней, поглощающих «лишние» нейтроны. Если же по мере выгорания топлива реактивность уменьшается, то управляющие стержни частично выводятся из системы и нейтронный поток вырастает до величины, необходимой для плановой работы реактора.

Характерное время отклонения от критического состояния определяется в основном периодом запаздывающих нейтронов, т. е. нейтронов, выделяющихся из осколков деления лишь некоторое время спустя после реакции распада. Этот период менее одной минуты, что предъявляет весьма жесткие требования к системе управления. Именно за это короткое время она должна «принять» и реализовать соответствующее решение при возникновении непредвиденных ситуаций.

Главная идея физически безопасного реактора — компоненты топлива должны быть подобраны так, чтобы, во-первых, его характерное время было заметно больше минуты и, во-вторых, чтобы в режиме его работы появились элементы саморегулирования.

Этого можно достичь, если в реакторе среди прочих реакций будет достаточно заметной следующая цепочка превращений:

$$^{238}\mathrm{U} + \mathrm{n} \longrightarrow {}^{239}\mathrm{U} \xrightarrow{\beta^-} {}^{239}\mathrm{Np} \xrightarrow{\beta^-} {}^{239}\mathrm{Pu},$$

где через ^{238}U, ^{239}U, ^{239}Np, ^{239}Pu обозначены соответствующие изотопы урана, нептуния и плутония, символом n нейтрон, символ β^- означает бета-распад (испускание электрона ядром). В этом случае образующийся в результате плутоний является основным и используемым сразу топливом. Характерное время такой реакции — время двух бета-распадов, равное приблизительно 2,5 суткам, т. е. оно почти на четыре порядка больше, чем для запаздывающих нейтронов.

Эффект саморегуляции связан с тем, что увеличение (по каким-то причинам) потока нейтронов приведет к быстрому выгоранию плутония, уменьшению его концентрации и соответственно потока нейтронов (образование же новых ядер ^{239}Pu будет идти в прежнем темпе примерно в течение 2,5 суток). Если же, наоборот, поток нейтронов в результате внешнего вмешательства резко уменьшится, то уменьшится скорость выгорания и увеличится темп наработки плутония с последующим увеличением числа выделяющихся в реакторе нейтронов через приблизительно такое же (равное нескольким суткам) время.

Описанные превращения протекают и в традиционных реакторах, однако в них они являются второстепенными для энерговыделения, поскольку используются в основном для накопления плутония. Ответ на вопрос, существует ли такая смесь U и Pu, при которой данная реакция становится доминирующей, может быть получен только с помощью математического моделирования этой сложнейшей системы.

Достаточно полная математическая модель активной зоны реактора должна включать в себя модели нестационарных пространственно трехмерных процессов переноса нейтронов в сильно неоднородной среде, выгорания топлив и реакторной кинетики, а также модель отвода тепла.

Однако для проверки осуществимости выдвинутой физической идеи, причем с хорошей количественной точностью, можно ограничиться более простой моделью. Первое упрощение — раздельный анализ нейтронно-ядерных процессов и процесса теплоотвода (оно оправдано при больших временах регулирования). Собственно нейтронные-

процессы вполне допустимо изучать не в трехмерной, а в одномерной геометрии рассматривая их к тому же в диффузионном и одногрупповом приближении (последнее означает, что соответствующим образом осредняются спектральные характеристики нейтронов).

Получающиеся при этих упрощениях уравнения переноса нейтронов по своему типу в некотором смысле близки к уравнениям теплопередачи (§ 2 гл. II). Они решаются совместно с уравнениями реакторной кинетики для шести групп предшественников запаздывающих нейтронов (уравнения типа радиоактивного распада, но с учетом «источников» нейтронов) и уравнениями выгорания для почти двадцати типов изотопов U, Pu, Np и других элементов. Разделение этих трех моделей невозможно, поскольку большинство искомых величин фигурирует во всех уравнениях. Метод решения — полудискретная (пространственная) аппроксимация исходной модели с последующим численным интегрированием по времени получающейся сложной системы обыкновенных дифференциальных уравнений высокой размерности.

Типичный вычислительный эксперимент состоит в задании начальной критической сборки (т. е. смеси веществ и геометрии активной зоны таких, чтобы получившийся реактор был критическим) и изучении эволюции со временем основных характеристик реактора. Заметим: начальная сборка находится с помощью специального предварительного вычислительного эксперимента.

Один из вариантов сборки в цилиндрической геометрии следующий:

Зона 1	Зона 2
^{239}Pu — 8% ^{238}U — 92% Сталь Na	Обедненный природный уран (^{235}U — 0,33%, ^{238}U — 99,77%) Сталь Na
$r_1 = 89{,}48$ см	$r_2 = 200$ см

Сборка разбита на две зоны разных размеров с разным содержанием плутония и изотопов урана (присутствие в гомогенной смеси натрия соответствует наличию теплоносителя, сталь отвечает конструкционным элементам). На рис. 83 приведена динамика (измеряемой в 10^{-6} с$^{-1}$) величины $\lambda(t)$ — характеристики времени отклонения от критичности для рассматриваемого двухзонного реактора (при $\lambda > 0$ число выделяющихся нейтронов пропорционально $e^{\lambda t}$). Видно, что в течение почти девяти месяцев работы у реактора без управления $|\lambda| \leqslant 10^{-6}$ с$^{-1}$, т. е. время его отклонения от критичности составляет примерно 10 суток (в данном варианте, если не вводится внешнее управление, реактор через девять месяцев затухает). Таким образом, убедительно подтверждается основная идея о физически безопасном реакторе.

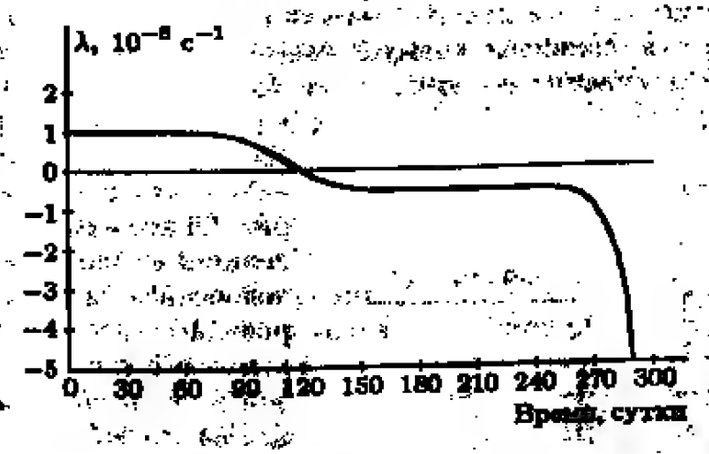


Рис. 83

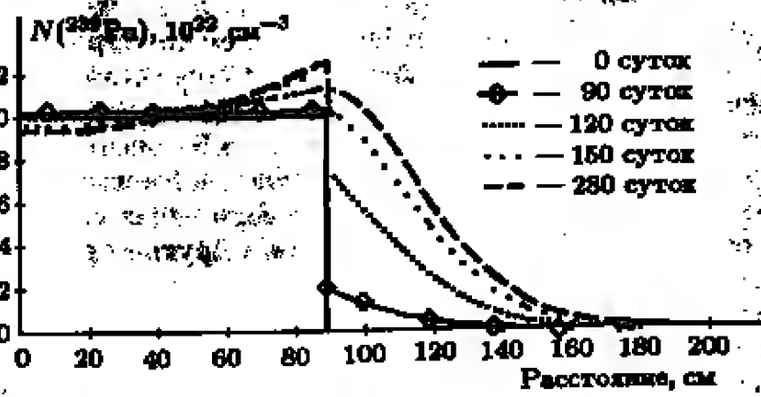


Рис. 84

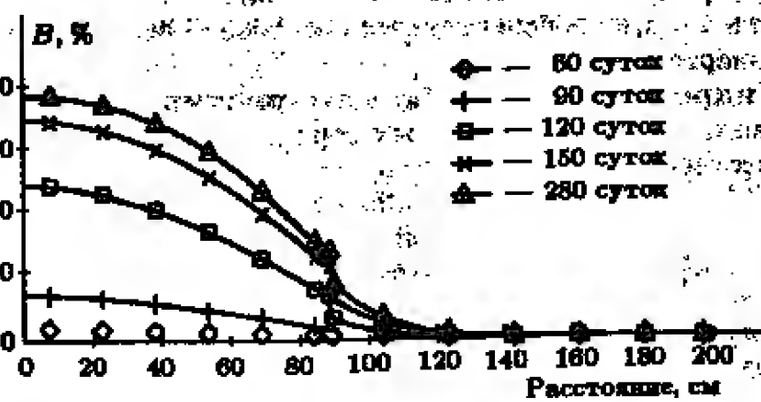


Рис. 85

Вычислительный эксперимент, кроме общего вывода, дает также ряд очень важных деталей работы изучаемой системы. На рис. 84 показана пространственная концентрация плутония $N(^{239}\mathrm{Pu})$ (в 10^{22} см$^{-3}$), а на рис. 85 — процентная степень выгорания топлива в различные моменты времени. В отличие от обычных реакторов здесь достигается высокая степень использования горючего (теоретически, без инженерных ограничений, 40–50% вместо стандартных 4–5%).

Эксперименты с моделью отвечают и на вопрос об устойчивости реактора к сильным возмущениям (авариям). Например, возможен выброс части теплоносителя (натрия) из активной зоны в результате самопроизвольного или насильственного разрыва оболочки трубопровода. На рис. 86 приведена зависимость величины λ от времени после подобной аварии. Она моделировалась следующим образом: в момент $t = 63$ сут. из первой зоны в интервале $0 \leqslant r \leqslant 15{,}4$ см за 15 мин с постоянной скоростью был выведен весь натрий. Это привело к возникновению колебаний характерного времени (при этом его минимальное значение составляло 8 мин, что вполне достаточно для реальной коррекции работы реактора). В дальнейшем возникшее возмущение гасится самим реактором без какого-либо внешнего воздействия (саморегулировка) и через 32 ч его характерное время стабилизируется, становясь равным, как и прежде, примерно 10 сут. Эксперименты с моделью обосновывают также и некоторые другие преимущества рассматриваемого реактора, например нет необходимости добавлять в активную зону новые порции плутония взамен выгоревшего (поэтому нет нужды в транспортировке и иных манипуляциях с одной из основных компонент ядерного оружия).

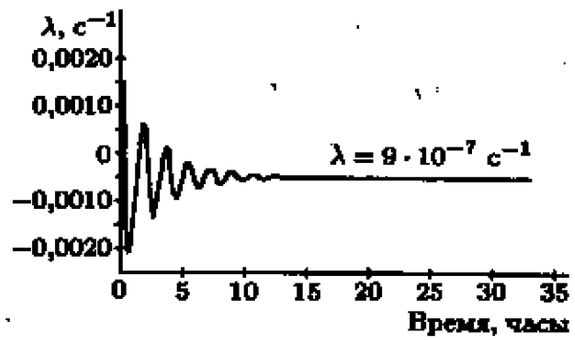


Рис. 86

Таким образом, математическое моделирование и вычислительный эксперимент позволяют не только качественно, но и количественно изучать один из перспективных способов повышения безопасности ядерной энергетики.

2. Гидрологический «барьер» против загрязнения грунтовых вод. Снабжение крупных городов и промышленных центров доброкачественной водой для питья и водой для технических нужд давно стало острой техно-экологической проблемой. Для ее решения помимо открытых и потому легко загрязняемых источников (реки, озера, водохранилища) активно используются подземные воды влагосодержащих пластов. Они менее подвержены антропогенным воздействиям, однако и для них вопросы, связанные с зачастую неизбежным загрязнением, остаются актуальными. Один из них — локализация вредных примесей, проникающих в часть пласта с тем, чтобы вода в других частях оставалась чистой и пригодной для потребления.

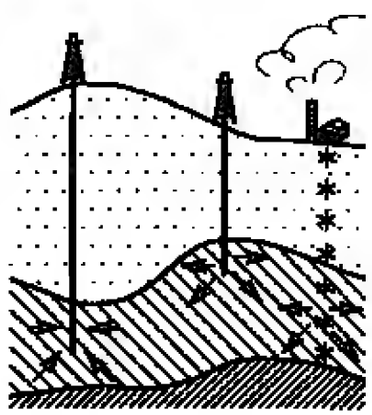


Рис. 87

Эту цель можно достичь, используя часть грунтовых вод для создания на пути распространения загрязнений своеобразного гидрологического барьера. Общая его схема показана на рис. 87: между источником загрязнения (звездочки) и водозаборными скважинами устанавливаются специальные скважины, накачивающие (достаточно чистую) воду в пласт и повышающие ее уровень (барьер). Накачка создает принудительное движение грунтовых вод вправо и влево от барьера (стрелки). Фильтрующаяся направо часть потока сносит назад текущую ей навстречу воду с примесями, препятствуя дальнейшему продвижению загрязнений вдоль пласта.

Математическая модель, реализующая эту схему, содержит уравнения фильтрации грунтовых вод и уравнения распространения примесей, дополненные соответствующими входными данными (свойствами грунта, воды и примесей, сведениями о геометрии рассматриваемой области, краевыми условиями и т. д.). При относительно небольшом содержании примесей в воде они не оказывают на ее движение заметного влияния, и поэтому фильтрацию можно рассматривать отдельно от динамики распространения загрязнений. При некоторых допущениях, главным из которых является предположение о достаточной протяженности пласта, движение воды описывается в рамках модели Буссинеска (§ 1 гл. II). В этом случае все характеристики процесса — функции переменных x, y, t (вид сверху на пласт изображен на рис. 88, где крестиками обозначены накачивающие, квадратиками — водозаборные скважины, а звездочками — источник загрязнений).

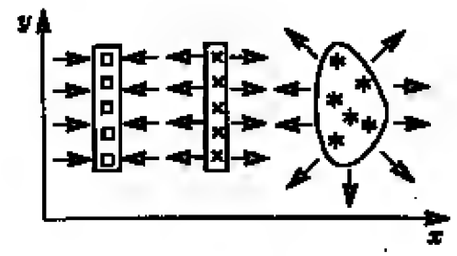


Рис. 88

Модель распространения примесей получается (как и уравнение Буссинеска) из закона сохранения (баланса) массы примеси в элементе грунта. В отсутствие диффузии загрязнений она представляет собой обычное уравнение неразрывности и в простейшем варианте имеет вид

$$\frac{\partial}{\partial t}[C(H+h)] + \frac{\partial}{\partial x}[C(H+h)u] + \frac{\partial}{\partial y}[C(H+h)v] = Q(x,y,t),$$

где $C(x,t)$ — искомая концентрация примесей, $Q(x,y,t)$ — известная интенсивность источников загрязнений (остальные обозначения те же, что и в § 1 гл. II). Это уравнение можно также трактовать (при известных функциях h, u, v) как двумерное уравнение переноса концентрации C. Так как по закону Дарси

$$u = -\nu\frac{\partial h}{\partial x}, \quad v = -\nu\frac{\partial h}{\partial y}, \qquad \nu = \mu\rho g,$$

то данное уравнение переписывается в форме

$$\frac{\partial}{\partial t}[C(H+h)] - \nu\frac{\partial}{\partial x}\left[C(H+h)\frac{\partial h}{\partial x}\right] - \nu\frac{\partial}{\partial y}\left[C(H+h)\frac{\partial h}{\partial y}\right] = Q(x,y,t).$$

Если же принять во внимание заметное хаотическое движение примесей, то последнее уравнение усложняется:

$$\frac{\partial}{\partial t}[C(H+h)] - \nu\frac{\partial}{\partial x}\left[C(H+h)\frac{\partial h}{\partial x}\right] - \nu\frac{\partial}{\partial y}\left[C(H+h)\frac{\partial h}{\partial y}\right] =$$

$$= \frac{\partial}{\partial x}\left[D(H+h)\frac{\partial C}{\partial x}\right] + \frac{\partial}{\partial y}\left[D(H+h)\frac{\partial C}{\partial y}\right] + Q(x,y,t),$$

становясь уравнением второго порядка (параболического типа) относительно функции C. Дополнительные члены в его правой части отвечают увеличению или уменьшению массы примеси в элементе грунта благодаря наличию потоков, вычисляемых по закону Фика

$$W = -D\,\mathrm{grad}\,C,$$

где $D > 0$ — коэффициент гидродинамической дисперсии (аналог коэффициента теплопроводности в законе Фурье; см. табл. 1).

Один из естественных способов перехода к дискретной модели — аппроксимация системы двух параболических уравнений: уравнения Буссинеска из § 1 гл. II для h и приведенного выше уравнения для C, разностной схемой, конструируемой, например, на основе интегро-интерполяционного метода (это особенно удобно в случае равномерной сетки). Заметим, что уравнение для h не содержит величины C, и это свойство модели значительно облегчает ее численное исследование. Разумно использовать безусловно устойчивые неявные схемы, позволяющие проводить расчет с достаточно крупным шагом по времени.

Получающаяся в итоге система нелинейных разностных уравнений решается следующим образом. Сначала из разностного уравнения Буссинеска для h_i^j (не содержащего функцию C_i^j) находятся значения h_i^j. Затем они используются в разностных уравнениях для C_i^j, откуда последние и определяются. При решении дискретных аналогов двумерных параболических уравнений для h и C применяется стандартный метод переменных направлений (и итерационные процедуры), заключающийся в последовательном решении прогонкой набора одномерных задач (сначала по направлению x, потом по направлению y, и наоборот).

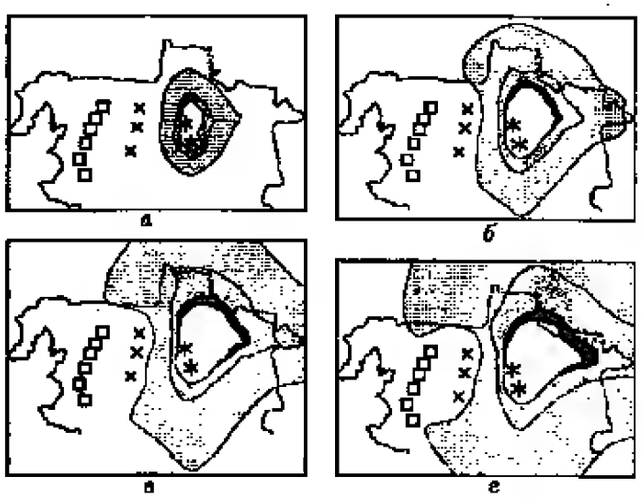


Рис. 89. а — $t = 2$ года; б — $t = 7$ лет; в — $t = 13$ лет; г — $t = 33$ года

На рис. 89, а–г приведены результаты демонстрационного вычислительного эксперимента по описанной выше методике, моделирующего динамику явления (замкнутые линии отвечают постоянным значениям концентрации примесей в некоторых условных единицах). Для простоты свойства грунта считались постоянными, на достаточно удаленном от источника загрязнений расстоянии величины h_i^j и C_i^j полагались равными нулю, расчетные шаги по времени и по пространству составляли примерно один месяц и один километр соответственно. Характерные масштабы всего процесса — десятки лет и сотни километров (размер небольшой центральноевропейской страны).

Из рис. 89 хорошо видно, что примеси от источника загрязнений распространяются во все стороны, кроме направления, где поставлен гидрологический барьер, предохраняющий заборные скважины от попадания в них недоброкачественной воды.

Разумеется, окончательный ответ о целесообразности гидрологического барьера может быть получен лишь после дополнительных исследований по моделированию изучаемых процессов, в том числе с учетом свойств конкретных грунтов, различных способов расположения скважин, экономических аспектов и т. д. Очевидно, что те или иные решения рассматриваемой проблемы должны опираться на математическое моделирование и вычислительный эксперимент, поскольку масштабы явления исключают натурные эксперименты, а лабораторные испытания в силу отсутствия подобия дают лишь ограниченную информацию.

3. Сложные режимы обтекания тел газом. Схожая с пп. 1, 2 ситуация характерна и для задач конструирования современной летательной техники. Аэродинамические трубы, в которых «продуваются» фрагменты будущих аппаратов, представляют собой очень дорогостоящие и сложные инженерные сооружения. Стоимость часа их эксплуатации может достигать нескольких тысяч долларов. Получаемые при этом экспериментальные данные не полны (ибо нельзя провести испытания при всех допустимых скоростях, углах атаки и т. д.) и требуют правильной расшифровки. К тому же они относятся лишь к отдельным частям аппаратов. Последний недостаток не может быть устранен по принципиальной причине. Картина обтекания газом помещенной в трубу маленькой материальной копии полного изделия имеет мало общего с процессами, происходящими на реальных (метры и десятки метров) масштабах, так как между ними нет детального подобия.

В этом легко убедиться, выписав два основных безразмерных параметра подобия (см. § 1 гл. V) для рассматриваемого класса явлений — число Маха ($\mathrm{M} = u/c$ — отношение скорости течения к скорости звука) и число Рейнольдса ($\mathrm{Re} = uL\rho/\mu$, где μ — коэффициент вязкости газа, L — характерный размер аппарата). Из вида чисел M и Re следует, что, сохраняя одинаковым для разных L число M, невозможно оставить неизменной также и величину $\mathrm{Re} = \mathrm{M}cL\rho/\mu$, характеризующую отношение сил давления к силам вязкости.

Поэтому основную информацию о поведении летательных аппаратов получают из вычислительных экспериментов, проводящихся во всем диапазоне допустимых параметров. Натурные же испытания служат в основном в качестве «эталонных» опытов: из сравнения их результатов с результатами математического моделирования выясня-

ется адекватность используемых моделей, точность вычислительных алгоритмов и при необходимости проводится их усовершенствование.

Продемонстрируем взаимоотношения вычислительного и натурного экспериментов на примере математического моделирования некоторых сложных течений с помощью методики, основанной на описанных в п. 6 § 4 гл. V кинетически согласованных разностных схемах для уравнений Эйлера и Навье—Стокса. В рассматриваемых ниже примерах процессы вязкого переноса существенны лишь в узкой области погранслоя около границы обтекаемых тел, поэтому решались уравнения Эйлера.

Одна из решавшихся эталонных задач — описание нестационарных двумерных течений в «колене» плоского прямоугольного канала, для которых имеются достаточно надежные экспериментальные данные. В нижнюю часть канала (рис. 90, 91), занимаемого вначале непо-

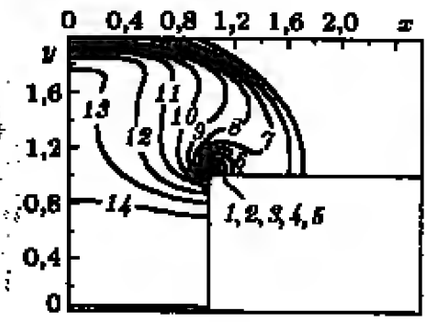


Рис. 90

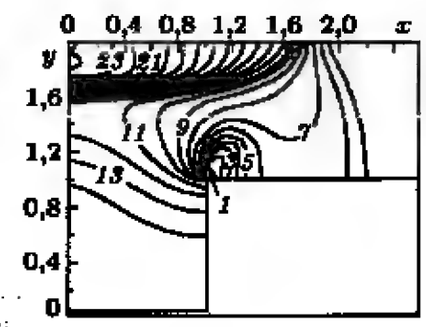


Рис. 91

движным газом ($\gamma = 1{,}4$), подается сверхзвуковой ($M = 1{,}2$) поток, при распространении которого вверх и направо последовательно возникает ряд сложных газодинамических конфигураций. Использовались сетки с числом узлов $N_x \cdot N_y \cong 3000$ с шагом по пространству $h_x = h_y = 1/27$ и по времени $\tau = 0{,}2$.

На рис. 90, 91 для моментов времени $t = 30$, $t = 40$ приведены эквидистантные с шагом $\delta\rho = 0{,}083$ изолинии плотности газа в канале единичного сечения ($\rho_1 = 0{,}747$, $\rho_{14} = 1{,}826$ для рис. 90 и $\rho_{24} = 2{,}656$ для рис. 91, единицы измерения условные). На рис. 90 хорошо видна зона уплотнения — ударная волна, образующаяся при огибании потоком прямоугольного выступа (линии сгущаются), и сформировавшаяся за ней зона разрежения. У верхней стенки канала газ тормозится и уплотняется. В результате этого торможения в течении появляется отразившаяся от стенки ударная волна (рис. 91, плотность газа в области уплотнения уменьшается при удалении от левой границы, что естественно).

В дальнейшем картина становится еще более сложной: отраженные от границ ударные волны распространяются по потоку, взаимодействуют между собой, вновь отражаются от стенок и т. д. Их динамика детально прослеживается в вычислительном эксперименте. Адекватность математического моделирования изучаемым явлениям подтверждается сопоставлением рис. 91 и соответствующего ему рис. 92, на котором представлена интерферограмма распределения плотности в потоке, полученная из прямого натурного эксперимента. Видно хо-

рошее качественное и количественное совпадение результатов вычислительного и реального экспериментов.

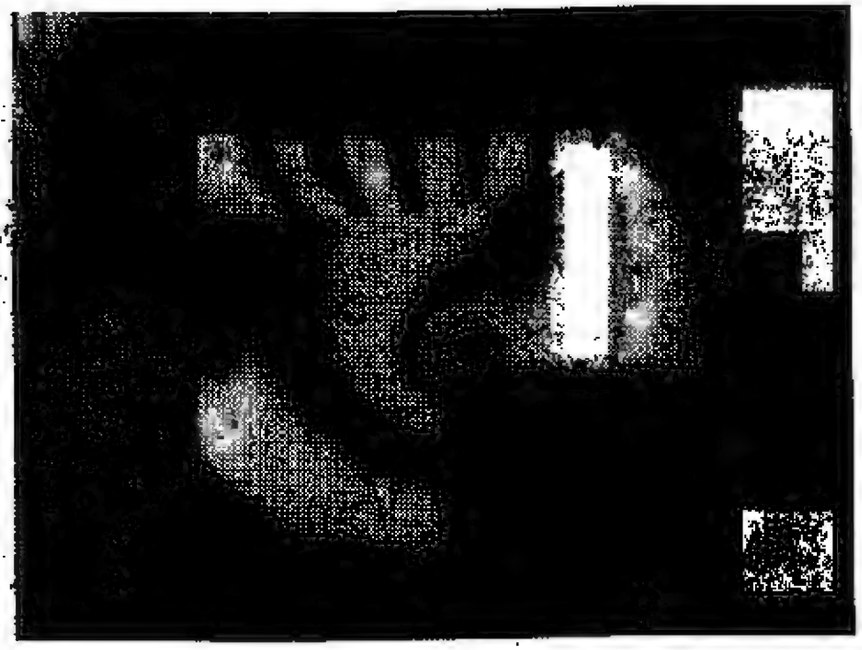


Рис. 92. Интерферограмма распределения плотности в натурном эксперименте

Таким образом, использовавшаяся численная модель вполне пригодна для описания течений сложной структуры, и с ее помощью можно изучать и другие процессы (в том числе те, для которых нет достоверных экспериментов). На рис. 93, 94 приведены результаты численного моделирования обтекания газом передней части цилиндрического тела с «иглой» — длинным тонким выступом, обращенным к набегающему справа (вдоль оси z) потоку. Параметры газа в потоке: $M = 3{,}5$, $\rho = 1$, $\gamma = 1{,}4$; размеры тела: диаметр основного цилиндра $D = 50$, диаметр иглы $d = 2$, ее длина $l = 46$. Расчет проводился в координатах r, z

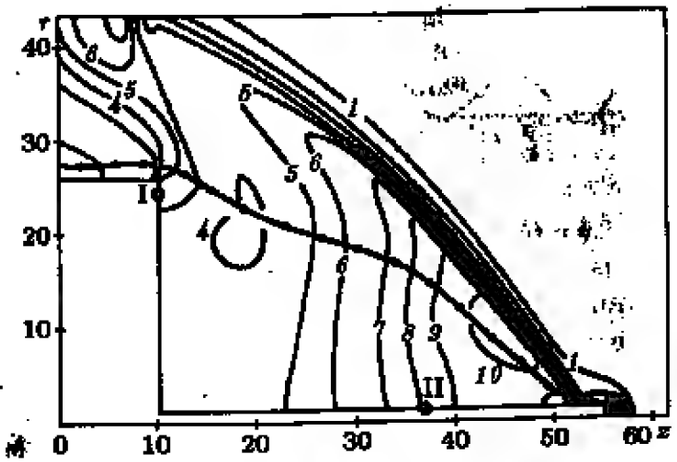


Рис. 93

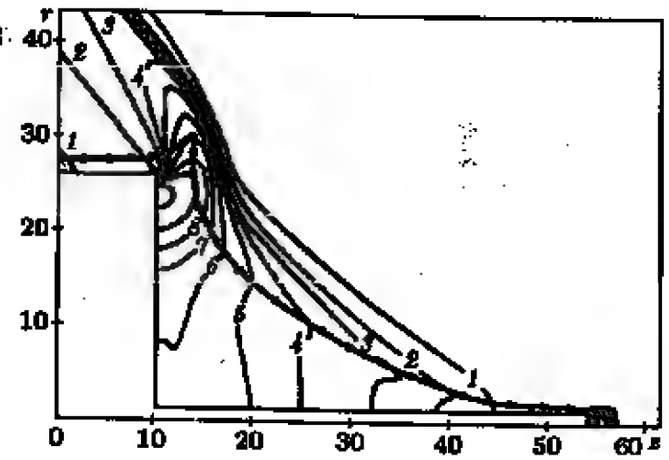


Рис. 94

на сетке, содержащей примерно 3000 узлов. Изображены эквидистантные изолинии давления (для рис. 93 $p_1 = 1{,}44$, $p_{10} = 11{,}47$, для рис. 94 $p_1 = 1{,}9$, $p_{10} = 24{,}44$), жирная линия с точками — звуковая линия с $M = 1$.

Сравнение рис. 93 ($t = 53$) и рис. 94 ($t = 83$) показывает, что обтекание имеет не установившийся, а сильно нестационарный характер (причем пульсации величин могут быть довольно значительными). Суть этих, как показывает численное моделирование, почти периодических колебаний состоит в увеличении и уменьшении размеров зоны течения между ударной волной (частые изолинии) и телом. Ударная волна, возникающая в результате торможения набегающего потока, перемещается со временем вдоль иглы почти по всей ее длине, что приводит к заметным изменениям параметров течения в окрестности иглы. На рис. 95 показана временная динамика давления в двух характерных точках тела (жирные точки на рис. 93). Линия *1* описывает давление в точке I, линия *2* — давление в точке II, линия *3* отвечает характерному размеру (по оси z) зоны между ударной волной и телом.

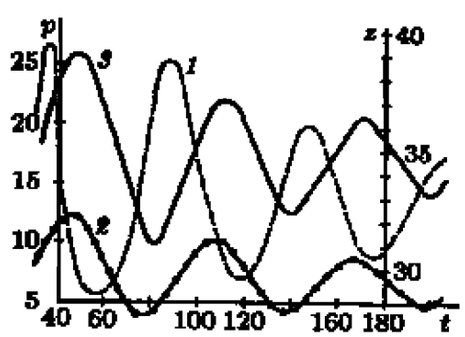


Рис. 95

Амплитуда колебаний весьма велика (в 2–3 раза больше среднего значения) так же, как и их частота, равная примерно l/c (где c — скорость звука в области между ударной волной и телом). Поэтому возникновение пульсаций (наблюдаемых также и в натурных экспериментах) и их количественные характеристики для различных параметров набегающих потоков и конфигураций тел обязательно необходимо учитывать при проектировании соответствующих летательных аппаратов. Подчеркнем, что постоянно возрастающая сложность подобных задач требует не только использования все более мощных компьютеров, но и (в неменьшей степени) постоянного совершенствования дискретных моделей и вычислительных алгоритмов.

4. Экологически приемлемые технологии сжигания углеводородных топлив. Большая часть используемой в промышленности и быту энергии производится в топочно-сжигающих устройствах — котлах ТЭЦ, газовых турбинах, двигателях внутреннего сгорания и т. д. Химическая энергия сжигаемых в воздухе углеводородов (бензин, керосин, мазут, метан) либо реализуется непосредственно для обогрева, либо преобразуется в механическую или электрическую энергию. Общее свойство всех этих процессов заключается в том, что воздух содержит не только кислород (окислитель), но и заметную долю «побочного» вещества — азота. С участием последнего в процессе горения образуются соединения, которые, в отличие от нейтрального азота, химически активны (главные из них — оксиды азота NO и NO_2). Вред от попадания оксидов в атмосферу разнообразен: они служат главным источником возникновения кислотных дождей, существенно способствуют истончению озонового слоя атмосферы, токсичны для дыхания и т. д.

Обычно используются два метода снижения выбросов оксидов азота. Первый из них — очистка продуктов горения. Соответствующее оборудование весьма эффективно, но одновременно и очень дорого. Стоимость современных фильтров достигает до 10–30% от стоимости самой ТЭЦ, в такой же пропорции увеличиваются расходы на ее текущую эксплуатацию. Основная идея второго метода — организация процесса горения таким способом, при котором образование оксидов азота было бы минимальным при сохранении энергетических характеристик установки. Осуществление этого подхода означало бы заметное снижение затрат на газоочистку продуктов горения от NO и NO_2 (полностью подавить их образование нельзя).

Происходящее в топочных устройствах процессы очень сложны. В них происходит множество химических превращений большого числа различных веществ, выделение и поглощение энергии, газодинамическое движение смесей, турбулентное перемешивание компонент горючего с воздухом и продуктами горения и т. д. Динамика этих явлений сильно зависит от режима подачи топлива, конфигурации котлов, расположения горелок и других характеристик топок. Поэтому поиск экологически допустимых режимов их работы не может опираться ни на чисто теоретические представления, ни на натурные эксперименты — дорогие, длительные и небезопасные.

Удовлетворительная по своей полноте математическая модель работы котла ТЭЦ включает в себя две взаимосвязанных части. Первая из них — локальное описание химической кинетики для 32 компонент смеси, в особенности тех реакций, которые играют наибольшую роль в образовании токсичных загрязнителей. Вторая часть — нестационарная пространственно двумерная модель макропроцессов тепломассопереноса (диффузия веществ, теплопередача, конвективное движение и т. д.). Она осложняется также необходимостью учета реальной геометрии котла (П-образная форма), наличием в нем нескольких ярусов горелок, служащих для впрыскивания топлива и воздуха, и другими факторами. Из вычислительных экспериментов с такой моделью находятся все необходимые характеристики горения при различных

режимах работы котла и отбираются оптимальные.

Охарактеризуем один из фрагментов математического моделирования описанных явлений, ограничившись первой частью — исследованием кинетики образования NO и NO_2 при горении CH_4 (метана), смешанного с воздухом (смесь считается изотермической и пространственно однородной). Полная модель включает в себя 196 прямых и столько же обратных реакций для 32 участвующих в горении веществ: O_2, N_2, CH_4, O, N, NO, NO_2 и т. д. В математическом отношении она представляет собой систему из $196 \times 2 = 392$ обыкновенных дифференциальных уравнений, описывающих прибыль и убыль компонент в результате соответствующих прямых и обратных реакций. Интегрирование ее по времени при известных начальных концентрациях веществ дает подробную динамику процесса. При этом прослеживаются не только суммарная концентрация того или иного вещества, но и все пути его образования, что особенно важно для анализа механизмов появления токсичных примесей (один из эффективных методических приемов, облегчающих расчеты: по результатам пробных вычислений отмечаются реакции, вносящие незначительный вклад в процесс и поэтому впоследствии в модели не учитываемые).

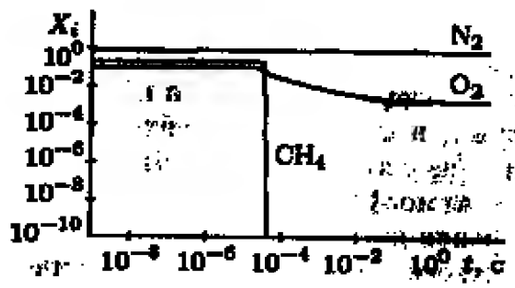


Рис. 96

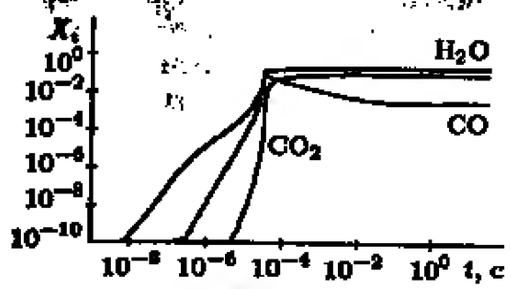


Рис. 97

Опишем один из типичных вычислительных экспериментов, проводившихся в близких к реальным условиям. Смесь находится при атмосферном давлении, ее температура равна 2000 К, а начальный состав в момент $t = 0$ задается в следующих долях: $[N_2] \approx 0{,}7$, $[O_2] \approx 0{,}2$, $[CH_4] \approx 0{,}1$. Расчет проводился до установления равновесия, т. е. до тех пор, когда концентрация любой из компонент перестает изменяться.

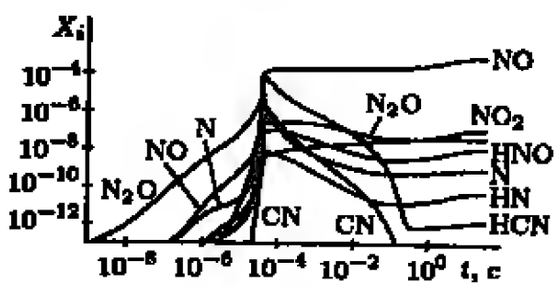


Рис. 98

На рис. 96 показана зависимость от времени для молярной (объемной) концентрации компонент исходной смеси, на рис. 97 — для H_2O, CO_2, CO, на рис. 98 — доли азотосодержащих веществ, в том числе NO, NO_2 и N_2O (закиси азота). Первый вывод — сгорание метана происходит очень быстро и его содержание вблизи момента $t = 4{,}1 \cdot 10^{-5}$ с падает практически до нуля. Процесс собственно горения после этого прекращается.

Рассмотрим теперь кинетику превращений азотосодержащих веществ. Прежде всего из рис. 98 видно, что итоговая концентрация NO на много порядков превышает концентрацию других вредных примесей. Поэтому основное внимание должно быть направлено на уменьшение именно этой компоненты. Существенно также то, что содержание NO монотонно увеличивается со временем. Значит, проблему уменьшения выхода NO проще решать, подавляя его образование (если же NO уже образовался, то обеспечить его исчезновение гораздо труднее).

Более важный результат следует из сопоставления динамики изменения содержания CH_4 и NO. Для NO эксперимент дает в моменты $t = 4 \cdot 10^{-5}$ с, $t = 0{,}1$ с, $t = 20$ с значения концентрации, равные соответственно $[NO] \approx 8 \cdot 10^{-5}$, $[NO] \approx 2 \cdot 10^{-4}$, $[NO] \approx 6 \cdot 10^{-4}$. Другими словами, к тому моменту, когда горение метана практически закончилось, в смеси присутствует лишь 13% от общего количества образующегося в итоге NO (основная часть вредных примесей появляется после того, как уже завершились энергетически полезные реакции сжигания метана).

Сильная временная разномасштабность этих двух процессов указывает на одну из принципиальных возможностей существенного уменьшения NO в продуктах горения. Необходимо как можно быстрее (в идеале сразу же после выгорания метана) удалить продукты горения из зоны пламени и резко охладить. В результате дальнейшие химические превращения, в том числе и образование NO, прекратятся, причем без потери энергетической мощности установки.

Моделирование ясно указывает и на ряд других важных свойств процесса, например, выход NO монотонно уменьшается при уменьшении температуры горения, выход закиси азота N_2O (озоноразрушающее вещество) заметен лишь на промежуточных стадиях процесса, в дальнейшем она превращается в безвредный N_2, и т. д. Эти выводы очень полезны для выработки рекомендаций по конструированию экологически приемлемых энергетических установок.

§ 2. Фундаментальные проблемы естествознания

Рассмотрим применение методологии математического моделирования к решению некоторых фундаментальных проблем из различных областей естествознания. Покажем, что она позволяет не только определять количественные характеристики изучаемых процессов, но и обнаруживать качественно новые явления.

1. Нелинейные эффекты в лазерной термоядерной плазме. Реальная перспектива решения энергетических проблем будущего связана с управляемым термоядерным синтезом (УТС) изотопов водорода, прежде всего дейтерия (D) и трития (T):

$$D + T = {}^4He + n + 17{,}6 \text{ МэВ}.$$

В этой элементарной реакции D + T образуются ядро гелия (α-частица) и нейтрон с суммарной кинетической энергией 17,6 МэВ, основная часть которой (14 МэВ) содержится в нейтронах (1 МэВ равен $1{,}602 \cdot 10^{-13}$ Дж). В грамме DT-смеси скрыта огромная энергия,

эквивалентная энергии, выделяющейся при сжигании 15 т угля. Несомненные преимущества УТС: практически неограниченные запасы в Мировом океане «горючего», обладающего колоссальной теплотворной способностью, а также экологическая чистота.

Для инициирования реакции D + T требуется нагреть смесь до нескольких десятков миллионов градусов и сжать ее до плотностей, сопоставимых хотя бы с плотностью DT-льда (охлажденной DT-смеси), равной 0,2 г/см^3. Если в дальнейшем удастся удержать ее в этом состоянии в течение времени, достаточного для того, чтобы заметная часть «топлива» выгорела, то выделится энергия, сравнимая с энергией, потраченной на нагрев и сжатие плазмы. Реализация этой схемы в лабораторных, а затем и в промышленных условиях означала бы практическое решение одной из старейших фундаментальных проблем физики.

О принципиальной осуществимости термоядерного синтеза известно с середины нынешнего века: в 40-х годах было установлено, что энергия излучения Солнца во многом обязана своим происхождением протекающим в его недрах реакциям слияния изотопов водорода. Впервые проведенные в 50-х годах взрывы водородных бомб — пример «рукотворной» термоядерной реакции.

Однако оба этих «способа» выделения энергии непригодны для целей УТС. На Солнце плазма удерживается в нужном состоянии за счет мощных гравитационных сил, а энергия, сопровождающая взрыв даже миниатюрной водородной бомбы, на много порядков превосходит приемлемую для мирного использования величину.

Существует несколько основных физических идей реализации УТС. Исторически первая из них — магнитное удержание, а наиболее разработанная конструкция — токамак, т. е. тороидальная камера, в которой внешние и генерируемые токами плазмы магнитные поля не дают разлететься и остыть нагретому плазменному «бублику». С этим относительно стационарным способом получения энергии синтеза конкурирует инерциальный синтез. Его идея состоит в том, чтобы быстро нагреть капельку горючего, которое из-за инерционности движения не успеет разлететься и охладиться, пока в ней не наработается нужное количество теромоядерной энергии. Периодически устраиваемые микровзрывы будут давать постоянные потоки нейтронов и α-частиц, утилизируемых за пределами рабочей камеры.

Наиболее подходящим источником столь быстрого нагрева являются лазеры. В лазерном термоядерном синтезе (ЛТС) последовательность событий такова. Излучение, направленное на сферическую мишень, поглощается в ее наружных слоях, нагревает и испаряет их. Образуется разлетающаяся с большими скоростями «корона», реактивная сила которой сжимает и нагревает ядро мишени. Далее происходит бурное термоядерное горение и выделение энергии, после чего мишень разлетается и остывает.

Первые пробные оценки показали, что для реализации ЛТС требуются лазеры, способные выделять излучение с энергией 10^8–10^9 Дж в течение нескольких наносекунд (1 нс = 10^{-9} с), фокусируя ее в крошечной области размером несколько десятков микрон. В настоящее

время лазеров с такими параметрами не существует, не предвидится их появления и в обозримом будущем.

Заметно снизить энергетический порог ЛТС позволяет использование различных нелинейных эффектов, присущих лазерной плазме. Они изучаются как экспериментально, так и теоретически с помощью вычислительного эксперимента с математическими моделями лазерных микровзрывов. Не характеризуя их подробно, поясним, что основой математического описания процессов в центральной части мишени служат уравнения теплопередачи и динамики сжимаемого газа (§ 2 и 4 гл. II). Для их дискретной аппроксимации успешно применяются полностью консервативные разностные схемы и схемы, получаемые из дискретных аналогов вариационных принципов (пп. 4, 5 § 4 гл. V).

Один из эффектов связан с типичной газодинамической нелинейностью. При сжатии ядра мишени в силу «градиентной катастрофы» (п. 7 § 4 гл. II) возникают ударные волны, которые нагревают ее центральную часть уже на начальной стадии процесса и мешают дальнейшему ее схлопыванию. В результате достигаются плотности гораздо меньшие, чем те, которые можно было бы получить при адиабатическом, безударном сжатии (и при той же энергии лазера). Это в свою очередь приводит к резкому уменьшению выхода термоядерной энергии, поскольку скорость реакции синтеза пропорциональна квадрату плотности горючего.

Однако процесс сжатия можно организовать так, чтобы он проходил безударным образом, когда все генерируемые сжимающим ядро «поршнем» газодинамические возмущения приходят в центр мишени одновременно и ударные волны не возникают вплоть до момента ее схлопывания. Одно из простых аналитических решений уравнений газовой динамики, демонстрирующее принципиальную осуществимость такого процесса, приведено в п. 3 § 3 гл. V.

Подробное математическое моделирование режимов безударного сверхсжатия мишеней убедительно показало, что оно действительно может быть реализовано и достигаемая в нем плотность вещества составляет десятки и сотни грамм на кубический сантиметр. При этом энергетический порог, необходимый для проведения «критического» опыта (вклад энергии в мишень и ее выход равны), снижается на несколько порядков и равен 10^5–10^6 Дж. Лазерный импульс, обеспечивающий сверхсжатие, должен быть «профилирован», т. е. его мощность должна изменяться со временем по закону $G(t) = G_0(t_f - t)^{-2}$ (режим с обострением; см. § 1–3 гл. V).

Схожего результата можно добиться, если заменить временное профилирование лазерного света пространственным (гидродинамическим) профилированием. Это достигается заметным усложнением конструкции мишени (см. рис. 99), но лазерный импульс как функция времени имеет традиционную «шапкообразную» форму.

Обе описанные концепции, детально проверенные в вычислительных экспериментах, имеют принципиальный характер для развития ЛТС, поскольку делают его одним из основных способов реализации управляемого синтеза. Они пока не могут быть проверены в прямых натурных экспериментах, поэтому адекватность математических мо-

делей ЛТС и надежность получаемых на их основе выводов устанавливается путем тщательного сравнения с существующими экспериментами по нагреву и сжатию лазерных мишеней.

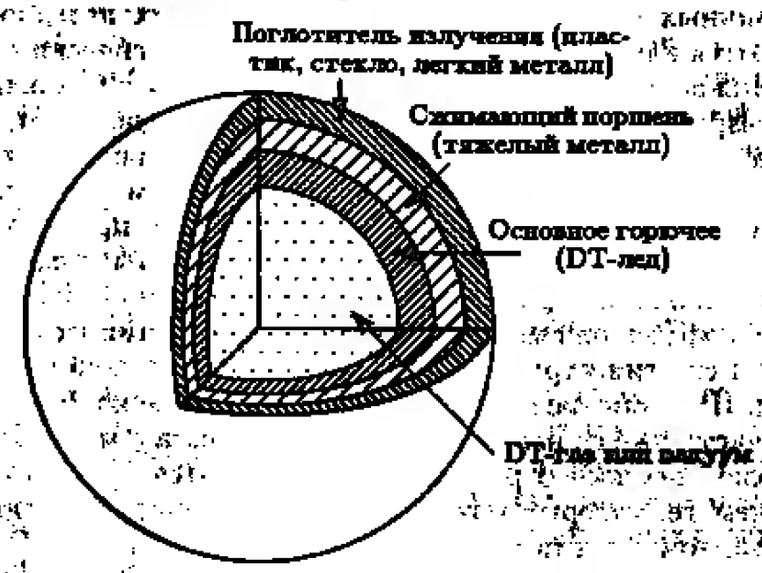


Рис. 99

Так, для мишеней сложной конструкции одной из центральных проблем является обеспечение симметрии и устойчивости сжатия. Толщина оболочек, окружающих DT-газ, составляет несколько процентов от начального размера мишени, и процесс сжатия можно уподобить превращению пустой яичной скорлупы в сгусток размером с маковое зернышко. Поэтому случающееся несовпадение результатов одномерных сферически симметричных расчетов сжатия с экспериментом неудивительно. В каждом таком случае требуется выяснение причин. На рис. 100 *а*, *б* приведена рассчитанная форма типичной экспериментальной мишени в момент наибольшего сжатия (*а* — общий вид, *б* — центральная часть; сгущение линий отвечает увеличению плотности). Видно, что ее центр несимметричен и, самое главное, сдвинут на 50 мкм относительно первоначального положения. В результате реальный нейтронный выход уменьшился в 100 раз в сравнении с рассчитанным по одномерной модели. Вычислительные эксперименты установили точную причину сдвига — несимметричность лазерного излучения, приходящего на поверхность мишени, что позволило выработать соответствующие практические рекомендации.

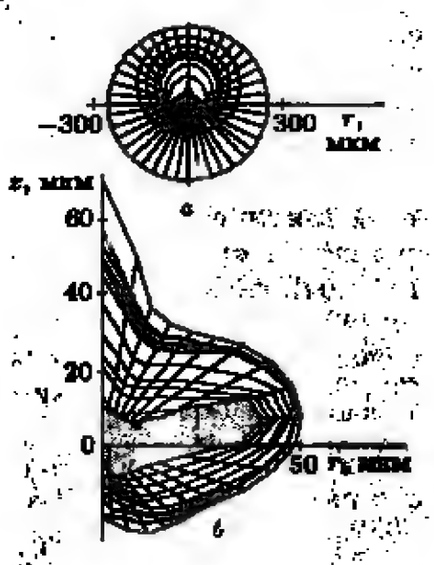


Рис. 100

Не менее важная причина потери одномерности сжатия — гидродинамическая неустойчивость на границе легкой и тяжелой жидко-

стей, испытывающих ускорение (благодаря ей ртуть, налитая на поверхность воды, находящейся в сосуде, и испытывающая действие сил тяжести, неизбежно перемешивается с водой, чего не происходит, если, наоборот, вода налита на поверхность ртути). Именно этой ситуации отвечает конечная стадия сжатия тяжелой оболочкой менее плотного центра: оболочка тормозится и поэтому ускорение, вызванное силами инерции, направлено в сторону от тяжелого к легкому веществу. Если неустойчивость успеет развиться достаточно сильно, то заметно нарушится симметричность сжатия и, более того, попадание части вещества оболочки в зону горения сразу же прекратит реакцию. Наблюдать этот тип неустойчивости непосредственно в экспериментах с мишенями нельзя. Поэтому используется универсальность математических моделей. Сначала сравниваются результаты моделирования с натурными экспериментами в ударных трубах — устройствах, где неустойчивость вызывается искусственно ударной волной, проходящей по разноплотной среде (это достигается разделением разных веществ диафрагмой). Затем делаются выводы для лазерных мишеней.

Рис. 101

На рис. 101 *а*, *б* показана форма границы раздела между тяжелым (ксенон) и легким (аргон) газами на момент $t = 100$ мс (длина волны начального возмущения $\lambda = 24$ мм); *а* — данные эксперимента, *б* — результаты расчетов. Видно, что численное моделирование дает не только типичную качественную картину («грибы»), но и вполне точно описывает количественные характеристики (амплитуду грибов и т. д.). В силу этого вычислительный эксперимент становится надежным инструментом изучения гидродинамической неустойчивости, в частности для определения допусков в точности изготовления мишеней.

Адекватность моделей лазерного синтеза позволяет рассматривать возможность использования еще одного характерного для плазмы типа нелинейности, благодаря которой в среде образуются локализованные структуры горения (см. § 3 гл. V). Если плазма достаточно плотная (несколько десятков граммов на кубический сантиметр), то α-частицы поглощаются в месте их выделения. Пусть, к тому же, в первом приближении гидродинамическим движением из-за его инерционности можно пренебречь в сравнении с процессами теплопередачи и горения. Тогда распространение выделяющейся энергии происходит в результате теплопередачи и описывается уравнением

$$\frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(k_0 T^{\sigma}\frac{\partial T}{\partial x}\right) + \frac{q_0 T^{\beta}}{1 + BT^{b}},$$

где $k_0 > 0$, $q_0 > 0$ — постоянные при коэффициенте теплопроводности плазмы и источнике энергии от термоядерных реакций. Для водородной плазмы $\sigma = 2{,}5$, а постоянные $\beta = 5{,}2$, $b = 3{,}6$ и B в источнике

таковы, что он может быть аппроксимирован степенной функцией вида $q_0T^{5,2}$ (при $1 < T < 3$–4 кэВ), $q_0T^{3,5}$ (при $T \approx 4$–5 кэВ) и $q_0T^{1,6}$ ($T > 5$ кэВ).

Следовательно, до температур примерно $5 \cdot 10^6$ К нелинейность среды отвечает локализованным LS- и S-режимам горения. На рис. 102 показана временная динамика структуры горения в плазме с плотностью 20 г/см^3 (численный расчет), на рис. 103 — такой же расчет, но

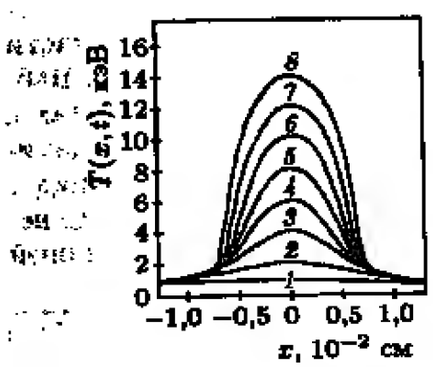


Рис. 102. 1 — $t = 0{,}0$ с; 2 — $t = 3{,}2 \cdot 10^{-10}$ с; 3 — $t = 3{,}8 \cdot 10^{-10}$ с; 4 — $t = 4{,}1 \cdot 10^{-10}$ с; 5 — $t = 4{,}3 \cdot 10^{-10}$ с; 6 — $t = 4{,}5 \cdot 10^{-10}$ с; 7 — $t = 4{,}7 \cdot 10^{-10}$ с

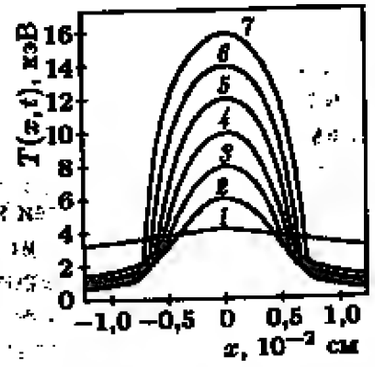


Рис. 103. 1 — $t = 0{,}0$ с; 2 — $t = 3{,}538 \cdot 10^{-8}$ с; 3 — $t = 3{,}715 \cdot 10^{-8}$ с; 4 — $t = 3{,}728 \cdot 10^{-8}$ с; 5 — $t = 3{,}732 \cdot 10^{-8}$ с; 6 — $t = 3{,}735 \cdot 10^{-8}$ с; 7 — $t = 3{,}737 \cdot 10^{-8}$ с; 8 — $t = 3{,}738 \cdot 10^{-8}$ с

при наличии собственного излучения из плазмы. Излучение даже несколько улучшает локализацию горения, которое не распространяется по среде и при более высоких, чем 5 кэВ, температурах. Поэтому при правильном инициировании термоядерной реакции возникшая в центре мишени область самолокализованного горения может послужить надежным «запалом» для «поджига» остальной массы горючего.

Стоимость современных экспериментальных установок для УТС — прототипов будущих электростанций — достигает сотен миллионов долларов, эксперименты на них сложны, трудоемки и длительны. Естественно, что при решении проблемы УТС ни одна физическая или технологическая идея всерьез не рассматривается без ее тщательного изучения методами математического моделирования и вычислительного эксперимента.

2. Математическая реставрация Тунгусского феномена. Наиболее яркий из известных примеров крупномасштабных столкновений небесного тела с атмосферой земли — Тунгусский феномен, наблюдавшийся около 7 ч 30 июня 1908 г. многими жителями обширного района Восточной Сибири с центром недалеко от поселка Ванавара (р. Подкаменная Тунгуска). Огромное светящееся космическое тело (угловой размер 0,5° на расстоянии 100 км, т. е. поперечный размер около 800 м) двигалось по ясному небу под некоторым углом к горизонту со скоростью более 1 км/с. Затем оно скрылось за лесом, после чего появилась яркая световая вспышка и многократные акустические волны, разбившие стекла окон в окрестности более 100 км. Свидетели ощущали также заметный тепловой импульс и наблюдали появление

теней, вызванных вспышкой. Геофизические и сейсмические станции России и всего мира зарегистрировали обошедшие несколько раз земной шар воздушные и сейсмические волны. Последующие экспедиции обнаружили сильные разрушения в тайге (вывал леса) и следы светового повреждения деревьев на площади около 2000 км^2 вокруг эпицентра события. Вещественных следов небесного тела найдено не было.

Все эти данные свидетельствуют — над тайгой произошел мощный воздушный взрыв (без образования кратера на поверхности Земли) с энергией не менее 10^6 т тротилового эквивалента.

Таким образом, сама природа поставила уникальный крупномасштабный эксперимент. Изучая его, можно получить многие важные сведения, относящиеся к астрономии, небесной механике, теории комет и метеороидов (так называют метеориты до их падения на Землю). Заметим, что метеориты (помимо образцов лунного грунта, доставленных с помощью космических экспедиций) являются единственными попадающими на Землю образцами вещества Вселенной.

Главное свойство Тунгусского явления: взрывной распад тела над Землей при отсутствии значительных количеств его вещества на ее поверхности говорит о том, что им не мог быть плотный каменный или железный метеороид. Им являлся либо редко встречающийся рыхлый каменный метеороид с повышенным содержанием льда, углерода и углеводородов, либо фрагмент ядра кометы — конгломерат кусков льда, газа и пыли. Все эти вещества легко испаряются или сгорают в атмосфере, не оставляя после себя следов. Кометная гипотеза полнее объясняет некоторые свойства явления. И если верна именно она, то Тунгусский феномен — единственный достоверный пример столкновения кометы с Землей (хотя вероятность такого события значительно меньше вероятности столкновения с рыхлыми метеороидом).

Общая гипотеза не дает ответа на вопрос об основных характеристиках космического тела — массе, скорости, угле падения, энергии, высвободившейся при взрыве, и т. д. Ее опровержение или подтверждение может быть получено лишь путем математического моделирования феномена и сопоставления результатов вычислительных экспериментов с имеющимися наблюдениями оставленных им разрушений.

Эта сложнейшая обратная задача математической реставрации события в общем виде формулируется следующим образом: в момент времени $t = 0$ в атмосфере на высоте $z = z_0$ задано движущееся со скоростью v_0 под углом θ_0 тело с линейными размерами L_0, плотностью ρ_0, температурой T_0, теплотой парообразования i_0 и характерной величиной напряжения на разрыв σ_0.

Начальное состояние атмосферы можно с хорошей точностью описывать в изотермическом приближении (температура постоянна) с плотностью и давлением, меняющимися с высотой по экспоненциальному закону: $\rho = \rho^0 \exp(-z/H_0)$, $p = p^0 \exp(-z/H_0)$, где H_0 — некоторая нормировочная постоянная.

В основу математического описания феномена кладутся нестационарные уравнения Навье—Стокса для сжимаемого вязкого теплопроводного газа (п. 5 § 3 гл. III), рассматриваемые в пространственно трехмерной геометрии. Поскольку явление характеризуется высокими

температурами и наличием излучения, то гидродинамические уравнения дополняются также трехмерным уравнением переноса излучения (§ 3 гл. II). Они должны решаться совместно при различных наборах перечисленных выше параметров, варьируемых в достаточно широких диапазонах. Для получения решения, естественно, необходимо задать свойства среды — коэффициенты вязкости, теплопроводности, поглощения света, уравнения состояния и т. д. В итоге задача содержит более десяти параметров, определяющих решение (применение методов подобия — см. § 1 гл. V — позволяет снизить их число и несколько упростить анализ результатов). Численная реализация описанных моделей осуществлялась с помощью соответствующих конечно-разностных схем.

Схематически течение событий, полученное из вычислительного эксперимента, изображено на рис. 104 (в данном случае $\theta_0 = 35°$, тело представляет собой полусферическую переднюю часть радиуса 70 м с примыкающим к ней цилиндром толщины 140 м). Штрихпунктирной линией обозначена траектория тела, сплошными — порождаемые его движением ударные волны в последовательные моменты времени. В моменты t_1, t_2 ударная волна баллистическая, аналогичная той, что появляется вокруг летящего со сверхзвуковой скоростью предмета, например истребителя. В момент t_3, соответствующий высоте тела $z = 7$ км, конфигурация ударной волны усложняется. К этому времени происходит сильное торможение и нагрев тела, в результате которого его вещество начинает взрывным образом расширяться, образуя сфероподобную сильную ударную волну. Дальнейшая динамика процесса, в том числе и характер разрушений, определяется обеими волнами. Они примерно одновременно (момент t_4) приходят к поверхности Земли и отражаются от нее (это справедливо лишь для правой, лежащей под траекторией части поверхности). Так как траектория наклонна, то картина не может быть симметричной относительно центра сферической ударной волны, принимаемой за эпицентр взрыва.

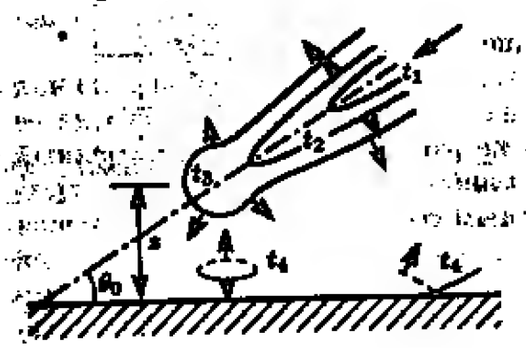


Рис. 104

Действительно, распространяющаяся против часовой стрелки верхняя часть баллистической ударной волны должна пройти до поверхности Земли, находящейся слева от эпицентра, большее расстояние, чем ее нижняя часть. Поэтому она не только достигает поверхности позже, но является также более слабой и производит менее значительные разрушения.

Вывод о несимметричности разрушений (на почти ровной поверхности тайги в месте катастрофы) блестяще совпадает с натурными измерениями. На рис. 105 показаны результаты обработки характеристик вывала леса в районе падения; круговые линии отвечают равным значениям силы разрушающих факторов, радиальные — направлению поваленных деревьев. Картина заметно несимметрична, она

имеет форму «бабочки», а не концентрических кругов, которые получились бы при вертикальном падении тела или при точечном взрыве космического аппарата с малой массой, не образующего из-за этого сильной баллистической ударной волны (высказывались гипотезы об искусственном происхождении Тунгусского тела).

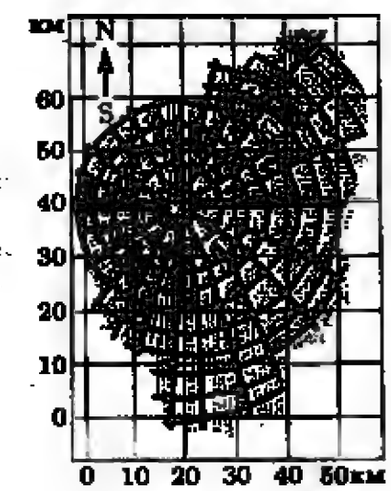


Рис. 105

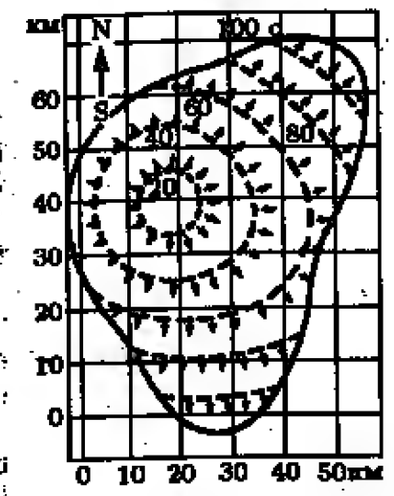


Рис. 106

На рис. 106 изображены аналогичные данные вычислительного эксперимента (стрелками обозначено направление движения воздуха по поверхности, штриховые линии отвечают положению фронта ударной волны, цифры около них — моменты времени в секундах), совпадающие не только качественно, но, при соответствующем подборе параметров тела, и количественно с последствиями феномена.

Относительно независимый способ проверки правильности выбора начальных характеристик тела дает моделирование действия излученного им теплового импульса (степени ожога деревьев). Сопоставление приведено на рис. 107. Точки обозначают данные натурных наблюдений: *1* — слабый ожог, *2* — умеренный, *3* — сильный (обугливание). Сплошная линия отвечает полученной из расчетов величине излучения $I = 16$ кал/см^2 в течение 2 с, необходимой для возгорания деревьев. Она практически точно совпадает с реальной границей зоны ожога.

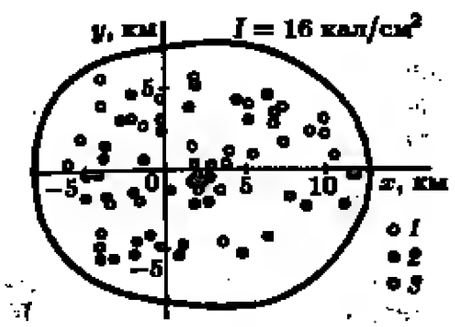


Рис. 107

Воспроизведение на основе математического моделирования Тунгусского феномена приводит к следующим вполне убедительным главным выводам, совокупность которых не может быть получена другими методами: тело общей массой 10^5 т вторглось в атмосферу под углом 35° со скоростью около 40 км/с, разрушилось, резко затормозилось на высоте $\sim 6{,}5$ км, ударные волны разрушили лесной массив, а излучение от нагретых до $10\text{–}12 \cdot 10^3$ К остатков тела произвело ожоги и воспламенение деревьев, мощность взрыва составляла примерно 15 Мт.

3. Климатические последствия ядерного конфликта. Достаточно точный долговременный прогноз погоды и изменений климата (вызванных в первую очередь антропогенными причинами) крайне необходимы для осмысленного планирования экономической, технологической, экологической и других видов активности человечества как в региональном, так и в глобальном масштабах. Оптимальное размещение индустриальных центров, приоритетное использование тех или иных видов сырьевых и энергетических ресурсов, предпочтительный выбор конкурирующих технологий, правильные акценты в агропромышленной политике — все эти проблемы находятся в тесной взаимозависимости с состоянием атмосферы, Мирового океана и поверхности Земли.

Протекающие в них геофизические процессы очень разнообразны и сложны. Они включают в себя гидродинамические движения атмосферного воздуха и вод морей и океанов, тепло- и массообмен в системе «океан—атмосфера», поглощение, рассеяние и отражение солнечного излучения (разного в разные времена года), сезонные изменения подстилающей поверхности и многие другие явления. Их сложность связана также с неоднородностью земной поверхности и внешней нестационарностью — вращением земного шара вокруг собственной оси и вокруг Солнца.

Неудивительно, что погодные и климатические явления отличаются сильной разномасштабностью во времени и в пространстве. Например, скорость ветра у данной точки поверхности Земли может во много раз отличаться от скоростей ветра на высотах над этой точкой. Кроме того, атмосферные и океанические потоки сильно турбулентны, т. е. характеризующие их величины испытывают хаотические флюктуации, масштаб которых достигает сотен и тысяч километров. Наконец, все упомянутые процессы существенно нелинейны и их отклик на изменение каких-либо параметров труднопредсказуем.

Поэтому прогноз погоды не может быть дан с достаточной гарантией. Этому препятствуют не столько ограниченность возможностей вычислительной техники, с помощью которой рассчитываются геометеополя, или отсутствие нужного числа данных, замеряемых стационарными и подвижными станциями, сколько принципиальные причины, связанные с масштабом и сложностью объекта. Точный прогноз погоды на несколько дней — вполне хороший результат.

Описанные трудности обязательно должны учитываться при оценке возможных последствий часто выдвигаемых многочисленных проектов, затрагивающих климатические процессы, таких, как, например, переброска части стока северных рек России в засушливые местности, перекрытие Берингова пролива, с тем, чтобы климат в регионе к югу от него стал более теплым, и т. д. Подобные эксперименты с системой, существующей в «единственном экземпляре», осуществимы лишь один раз, их результаты необратимы и должны быть известны с научно обоснованной точностью. Поэтому основным средством анализа и прогноза данных объектов служит вычислительный эксперимент с их математическими моделями.

Некоторые проблемы математического моделирования этих явлений облегчаются, если речь идет о долгосрочных изменениях погоды (месяцы) и климата (годы, десятилетия). Значения средних величин на достаточно длинном промежутке времени в меньшей степени зависят от мелкомасштабных пульсаций и неустойчивостей. Климатические модели системы «атмосфера—океан» состоят из ряда взаимосвязанных блоков: трехмерные нестационарные уравнения движения сжимаемой и несжимаемой жидкости с учетом вязкости и теплопроводности (типа уравнений Навье—Стокса), уравнения переноса излучения в атмосфере и т. д. Они решаются совместно при заданном начальном состоянии системы (в том числе состоянии подстилающей поверхности) и известной динамике солнечной радиации и других внешних факторов. Их дискретные аналоги, представляют собой, как правило, разностные схемы. Так как вычислительные алгоритмы для решения этих проблем с необходимостью должны быть экономичными и обладать хорошей разрешающей способностью (точностью), то при построении разностных схем существенно используются, например, описанные в § 4 гл. V подходы.

Климатические вычислительные эксперименты условно можно подразделить на два типа. К первому относятся те, что призваны удостоверить адекватность моделей путем сопоставления их результатов с имеющимися сейчас надежными данными натурных наблюдений. Эксперименты второго рода направлены на прогноз долгосрочных климатических изменений, вызванных естественными или искусственными причинами.

На рис. 108 дано сопоставление «калибровочных» вычислительных экспериментов по одной из наиболее полных моделей климата с данными наблюдений. Измеряемая величина $\langle v \rangle$ — средняя за год скорость ветра (в м/с) для всевозможных широт на высоте, отвечающей давлениям 400 мбар и 800 мбар (кривые *1* и *2*). Штриховые линии — результаты наблюдений. Расчеты проводились на установление: задавались типичные для данного времени года внешние воздействия и счет велся до выхода процесса на квазистационарный режим, соответствующий выбранному месяцу, после чего величины осреднялись по времени и по пространству. Соответствие результатов вычислительных экспериментов и наблюдений оказалось вполне удовлетворительным для столь сложного объекта, особенно в зоне экватора (пассатные ветры).

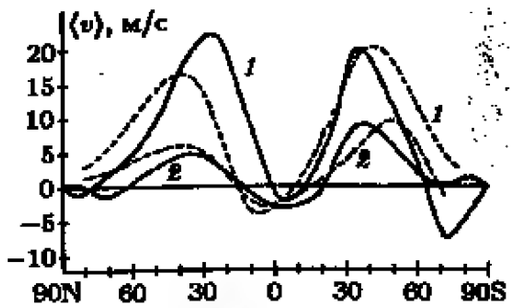


Рис. 108

Установленная адекватность климатической модели дает возможность проводить прогностические эксперименты, например связанные с проблемой «парникового эффекта». Большое количество CO_2 (окиси углерода) антропогенного происхождения выбрасывается в воздух и, препятствуя выходу собственного излучения Земли в космос, способствует увеличению средней температуры атмосферы. Это в свою

Таблица 4

	1	2	3
Средняя температура атмосферы, °С	−19,2	−17,54	−17,0
Температура воздуха у подстилающей поверхности, °С	13,9	15,3	15,8
Температура подстилающей поверхности, °С	12,5	13,6	14,3
Поток коротковолновой радиации на подстилающую поверхность, Вт/м^2	255,3	254,3	254,8
Осадки, мм/сут	2,04	2,15	2,17

очередь может привести к интенсивному таянию льдов, повышению уровня Мирового Океана и другим негативным глобальным последствиям. В табл. 4 приведены некоторые данные математического моделирования системы «атмосфера—океан» при различных концентрациях CO_2 в воздухе. Первый столбец отвечает современной концентрации, второй и третий — соответственно удвоенному и учетверенному содержанию CO_2.

Вычислительный эксперимент для этих двух гипотетических ситуаций проводился до установления нового, отличного от нынешнего квазиравновесия в системе с последующим осреднением результатов. Видно, что средняя температура как воздуха, так и подстилающей поверхности заметно увеличивается (изменение глобальной температуры на 1–2°С считается значительным), растет и среднее число осадков. Эксперимент также хорошо демонстрирует нелинейность объекта. Увеличение средней температуры при переходе от удвоенного к учетверенному содержанию CO_2 намного меньше, чем при удвоении современной концентрации. Отклик системы «атмосфера—океан» на внешнее (в данном случае антропогенное) воздействие не пропорционален величине этого воздействия (в рассматриваемой ситуации система смягчает последствия увеличения индустриальной активности человека).

Математическое моделирование позволяет оценивать результаты не только плавных, как в описанном выше сценарии, но и резких внешних вмешательств в систему. Одним из них может быть ядерный конфликт между воюющими державами. Полная неприемлемость для цивилизации глобальной ядерной войны давно общепризнана. Однако в свое время рассматривалась возможность ограниченного обмена ударами («атака на города») с использованием малой части боезапаса. Каковы будут климатические последствия ограниченной ядерной войны?

Опыт массированных бомбардировок крупных городов во время Второй мировой войны свидетельствует о неизбежности возникновения в них огромных пожаров. Их интенсивность такова, что в огне сгорают не только легковоспламеняющиеся материалы (дерево, пластмассы), но и негорючие в обычных условиях асфальт, бетон, кирпич. В отличие от относительно чистого горения лесов, мощные городские пожары будут сопровождаться выбросом в атмосферу огромного количества сажи — по некоторым оценкам, примерно по 1 т сажи на 1 т тротилового эквивалента заряда. Это значит, что ядерная атака городов с суммарной мощностью 100 Мт (примерно 1% от общего боезапаса ядерных держав) приведет к немедленному попаданию в атмосферу 10^8 т сажи.

Такая степень «задымления» в несколько десятков раз уменьшит поток солнечного света у подстилающей поверхности. Вычислительные эксперименты имитировали именно этот сценарий: в моделях мгновенно изменялись соответствующие характеристики атмосферы над наиболее вероятными районами возможного конфликта и прослеживалась временная динамика климатических величин.

Главным эффектом является быстрое и исключительно сильное охлаждение воздуха над континентами: даже в случае использования всего 1% имеющегося в наличии боезапаса средняя температура у подстилающей поверхности через неделю упадет на 15°С. Средняя температура более высоких слоев атмосферы, наоборот, увеличится примерно на такую же величину (поскольку в них поглощается вся солнечная радиация). Образующаяся температурная инверсия чрезвычайно стабильна («холодное» — внизу, «теплое» — вверху) и сохранится в течение многих месяцев.

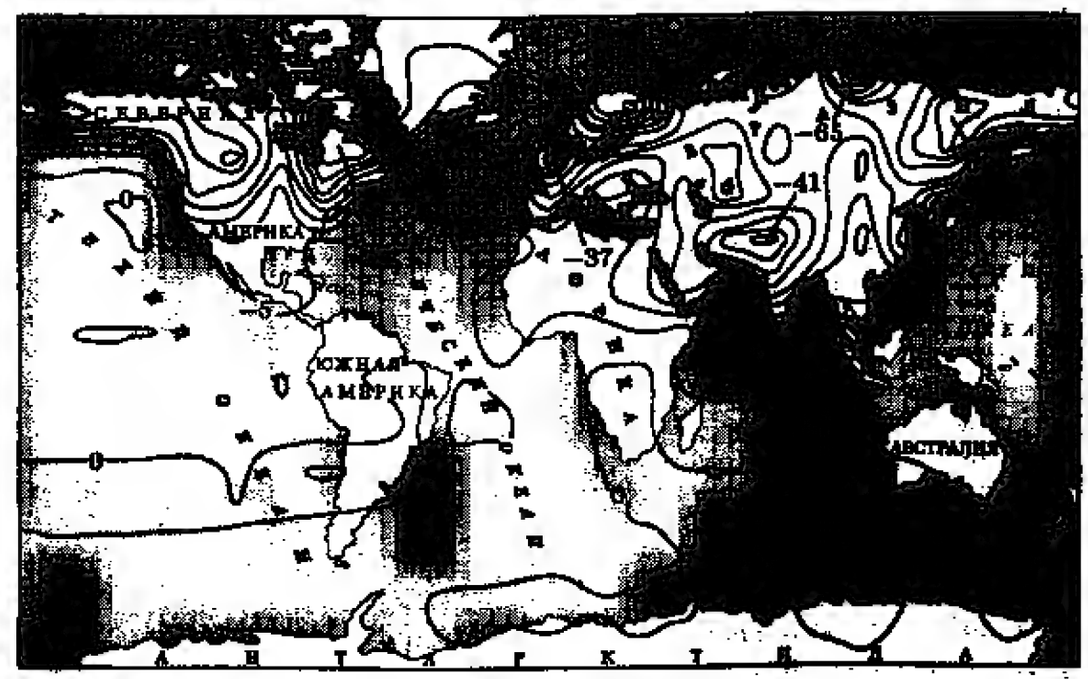


Рис. 109

Подобное развитие событий нельзя назвать иначе как глобальной климатической катастрофой. При этом средние климатические величины не дают о ней полного представления. На рис. 109 показаны изолинии температуры воздуха у поверхности Земли на 30–40 день после «100-мегатонного конфликта». Температура упадет ниже нормы на 56°С на севере Европы, на 65°С на севере Сибири, на 43°С в

Северной Америке и на 41°С на юге Азии и т. д. На высоте горных ледников температура станет намного выше нормальной, что приведет к бурным паводкам. Огромные массы воды, попав на переохлажденные равнины, покроют их ледяной коркой. Океан из-за своей большой теплоемкости будет остывать гораздо медленнее, и контраст температур между водой и сушей породит невиданной силы ураганы в обширных прибрежных районах.

Катастрофа действительно глобальна еще и потому, что дым, выделившийся в отдельном регионе, распространится на всю планету и принесет «ядерную зиму» в любую точку, не исключая и точки, из которой был нанесен внезапный односторонний удар.

Ранее считалось, что основными поражающими факторами ядерного оружия являются проникающая радиация и ударные волны. Математическое моделирование убедительно свидетельствует: помимо этих (относительно локальных) факторов ядерный конфликт будет сопровождаться глобальными катастрофическими изменениями климата, и поэтому он неприемлем даже в ограниченном варианте.

4. Магнитогидродинамическое «динамо» Солнца. Двадцатидвухлетние циклы солнечной активности имеют немаловажное значение для земной жизнедеятельности. Заметное увеличение амплитуды магнитного поля, генерируемого Солнцем, сказывается на числе заболеваний, на условиях сельскохозяйственного производства, на надежности функционирования средств радио- и телесвязи.

Качественное объяснение и количественное описание солнечной «магнитной синусоиды» относится к числу фундаментальных естественнонаучных проблем, содержащих, к тому же, парадоксальные загадки. Например, как свидетельствуют измерения ежегодных приростов древесной массы (толщины колец на срезах деревьев), несколько веков назад колебания магнитного поля Солнца прекращались более чем на 50 лет.

Высказывались различные соображения о природе этого явления, в том числе и о влиянии Юпитера, период обращения которого вокруг Солнца составляет 11,9 лет. В середине 50-х годов была сформулирована первая, претендующая на научную достоверность концепция — теория солнечного МГД-динамо.

Под этим термином понимается сложная совокупность процессов, служащих причиной периодической генерации магнитных полей светила. Они включают в себя конвективные и турбулентные гидродинамические движения солнечной плазмы на его поверхности и во внутренних слоях. Движущиеся заряженные частицы (токи) создают магнитное поле, преобразуя в него часть своей кинетической энергии. В свою очередь возникшее магнитное поле влияет на движение электронов и ионов и тем самым создаются предпосылки для появления в системе периодического процесса.

Проверка справедливости теории МГД-динамо может быть осуществлена лишь путем вычислительных экспериментов с математическими моделями генерации солнечных магнитных полей и сопоставления (там, где это удается) их результатов с данными наблюдений. Основные уравнения этих моделей получаются при соответствующих

предположениях из системы уравнений Максвелла и имеют вид

$$\frac{\partial \vec{B}}{\partial t} = \operatorname{rot}[\vec{B} \times [\vec{r} \times \vec{\omega}]] + \operatorname{rot} \alpha \vec{B} - \operatorname{rot} \beta \operatorname{rot} \frac{\vec{B}}{\mu},$$

$$\operatorname{div} \vec{B} = 0.$$

Здесь $\vec{B} = \vec{B}(\vec{r}, t)$ — искомый вектор магнитной индукции, $\vec{r}$ и t — радиус-вектор и время, $\vec{\omega} = \vec{\omega}(\vec{r})$ — заданный вектор угловой скорости, зависящий от пространственных координат (поскольку Солнце, не будучи твердым телом, вращается отнюдь не как единое целое), $\alpha = \alpha(\vec{r})$, $\beta = \beta(\vec{r})$, $\mu = \mu(\vec{r})$ — задаваемые усредненные характеристики соответственно турбулентно-конвективного движения, турбулентной проводимости и магнитной проницаемости плазмы. Заметим, что входные данные $\vec{\omega}$, α, β, μ точно неизвестны, и поэтому адекватность результатов моделирования должна проверяться особенно тщательно.

Решение проблемы несколько облегчается тем, что основная 22-летняя составляющая крупномасштабного колеблющегося магнитного поля Солнца является осесимметричной (не зависит от азимутального, меридионального угла). Вместо исходной трехмерной задачи можно ограничиться анализом процесса в двумерной постановке.

Однако и при таком заметном упрощении модель все равно остается достаточно сложной. Требуется, варьируя входные данные, численно решать большое количество нестационарных задач для системы двух квазилинейных параболических уравнений в области $r < R$, где R — радиус Солнца (внутренние задачи), и столько же задач для эллиптического уравнения в области $r > R$, где $\beta = \infty$ (внешние задачи). При этом, разумеется, их решения должны удовлетворять некоторым условиям сопряжения при $r = R$. Дополнительная сложность, сопровождающая решение внутренних задач, — наличие в области $r < R$ двух зон: лучистого ядра $0 < r < R_0 < R$ (где величина β мала, но пренебречь ею нельзя) и конвективной зоны $R_0 < r < R$, на границе которой с ядром сильно изменяется не только β, но и функция μ.

Дискретный аналог рассматриваемой модели основан на разностной аппроксимации дифференциальных уравнений для внутренней задачи и интегрального уравнения для внешней задачи (в области $r > R$ удобнее численно решать интегральное уравнение, эквивалентное исходному эллиптическому). Для численного интегрирования системы параболических уравнений применялись неявные абсолютно устойчивые разностные схемы переменных направлений, решаемые различными вариантами метода прогонки. Зная значения сеточных функций на границе $r = R$, относительно нетрудно, используя условия сопряжения, найти решение разностного интегрального уравнения при $r > R$ и тем самым полностью решить задачу. При построении дискретных аппроксимаций учитывалась возможность существования резких изменений коэффициентов уравнений по пространству (см. п. 3 § 4 гл. V). Также было обеспечено свойство симметрии численного решения (четность или нечетность) относительного солнечного экватора, присущее решению исходной задачи при симметричных начальных данных.

Вычислительные эксперименты показали, что качественный и во многом количественный характер процесса зависит от величины основного безразмерного параметра задачи — числа $D = R^3\alpha_0\omega_0/\beta_0^2$, где $\alpha_0, \omega_0, \beta_0$ — некоторые характерные значения функций α, ω, β. При $D \approx D_{кр}$ колебания происходят с постоянным периодом и постоянной амплитудой. При $D > D_{кр}$ $(D < D_{кр})$ колебания имеют несколько иной период и нарастают (затухают).

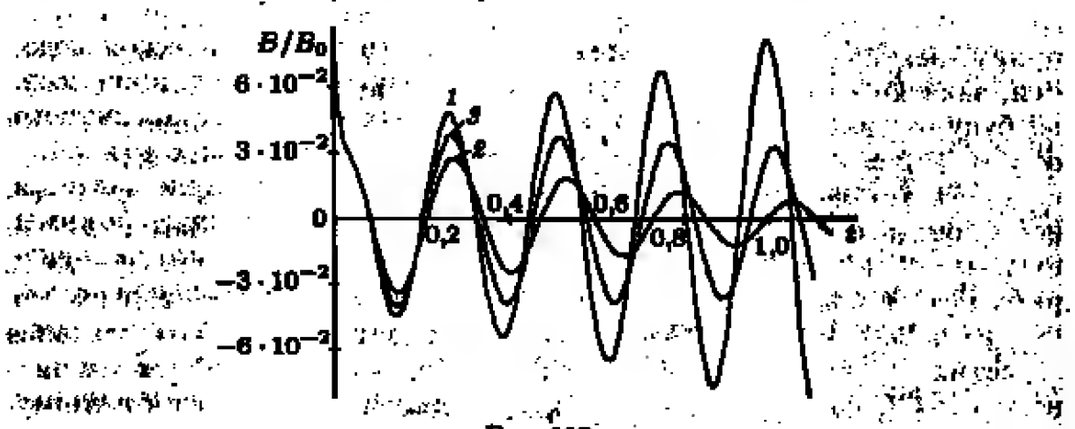


Рис. 110

Элементарная интерпретация этого результата следует из анализа структуры числа D. Пусть, например, величины R, α_0, β_0 постоянны, а ω_0 изменяется. При малых ω_0 (число D также мало) энергии вращения Солнца недостаточно для поддержания периодического процесса и он, стартовав, со временем затухает. При больших ω_0 солнечная «динамомашина» работает слишком интенсивно, генерируя растущее по амплитуде поле. Наконец, если комбинация характерных параметров отвечает критическому значению $D_{кр}$, то колебания носят наблюдаемый в действительности регулярный характер.

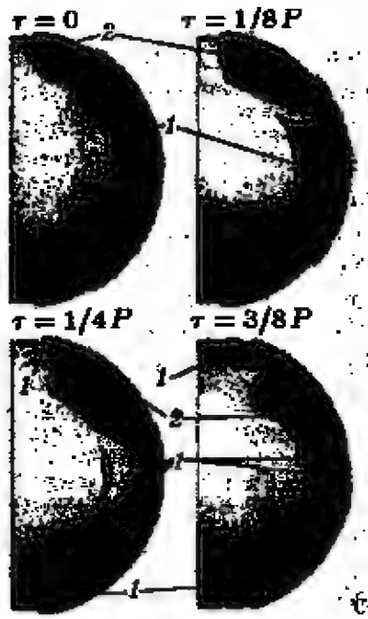


Рис. 111

На рис. 110 изображены (в безразмерных единицах) результаты вычислительных экспериментов, соответствующих описанным режимам (кривая *1* отвечает случаю $D = 2{,}75 \cdot 10^4 > D_{кр}$, кривая *2* — $D = 2{,}3 \cdot 10^4 < D_{кр}$, кривая *3* соответствует значению $D = D_{кр} = 2{,}56 \cdot 10^4$ и регулярным колебаниям с периодом $P = 0{,}248$). Значения периода и амплитуды колебаний в установившемся режиме не зависят от принимаемых в расчетах начальных распределений и с хорошей количественной точностью согласуются с данными наблюдений.

Данный вывод не связан с «подгонкой» неточно известных функций $\vec{\omega}$, α, β, μ для получения из модели заранее известного результата. В этом убеждает сопоставление более тонких эффектов, сопровождающих солнечную активность. На рис. 111 приведены временные баттерфляй-диаграммы — имеющая вид крыла бабочки конфигурация линий равного уровня магнитного поля на поверхности Солнца (расчетные результаты за половину периода). Зоны с высокой величиной магнитного поля (цифра *2* — им отвечают солнечные «пятна») со временем смещаются из полярных областей к экватору, а их место занимают зоны с меньшим магнитным полем (цифра *1*), и наоборот. Процесс периодически повторяется, хорошо согласуясь с реальной картиной. Еще один из способов подтверждения адекватности математической модели солнечного «динамо» — переход к анализу трехмерных процессов. Они описывают динамику неосесимметричных компонент магнитного поля Солнца, совпадающую во многих чертах с реальной картиной.

Математическое моделирование показывает, что Солнце является нелинейным МГД-генератором, допускающим неожиданные режимы работы (именно с нелинейностью связываются паузы в его активности). Поэтому их знание важно не только для фундаментальной науки, но и для практических целей.

§ 3. Вычислительный эксперимент с моделями трудноформализуемых объектов

Продемонстрируем универсальность математического моделирования на примерах исследования трудноформализуемых объектов, для которых не существуют или неизвестны точно сформулированные законы. Покажем, что его применение дает разнообразные возможности для более глубокого понимания их принципиальных свойств.

1. Диссипативные биологические структуры. В рассматриваемой в п. 1 § 3 гл. IV биологической модели типа «хищник—жертва» игнорируется возможность неоднородного размещения популяции в занимаемой ею части пространства. Такие модели служат лишь первым приближением к реальности. В действительности условия проживания популяции никогда не бывают одинаковыми в разных частях ареала. Кроме того, даже для пространственно однородной среды обитания всегда существенны чисто биологические причины скопления или разрежения представителей популяции: инстинктивные поведенческие мотивы собирания их в стаи и стада, сезонные изменения в природе (например, наступление поры нереста или взросления птенцов) и т. д.

Поэтому более подробные математические описания популяций должны учитывать пространственные факторы. Одна из типичных биологических моделей такого рода выглядит так:

$$\frac{dN}{dt} = (\alpha - cM)\,N + D_N\,\frac{\partial^2 N}{\partial x^2},$$

$$\frac{dM}{dt} = (-\beta + \gamma N)\,M + D_M\,\frac{\partial^2 M}{\partial x^2},$$

где t — время, x — пространственная координата (для простоты пространство считается одномерным), $N(x,t)$ и $M(x,t)$ — «плотности» жертв

и хищников, $\alpha > 0$, $c > 0$, $\beta > 0$, $\gamma > 0$, $D_N > 0$, $D_M > 0$ — постоянные, характеризующие внутренние свойства популяций.

От уравнений Лотки—Вольтерра данная модель отличается присутствием в правой части членов «диффузионного» вида (D_N, D_M — коэффициенты «диффузии») и представляет собой систему двух уравнений параболического типа относительно величин N и M. Происхождение этих «диссипативных» членов обосновывается теми же предположениями, что делались в п. 1 § 1 гл. IV при получении модели динамики скопления амеб: на скорость изменения численности популяции влияет наличие «хаотического» движения особей в пространстве, формирующее их поток из более «заселенных» в менее «заселенные» местности (он считается пропорциональным градиенту их плотности).

Поведение численности популяции в пространственных моделях может кардинальным образом отличаться от картины, даваемой точечными моделями. Рассмотрим, например, такой эффект, как формирование в пространстве «волны погони» хищника за жертвой. На рис. 112 показаны результаты демонстрационного вычислительного эксперимента с описанной моделью (единицы измерения условные). Задача рассматривалась в неограниченной области (задача Коши), начальная плотность жертв — экспоненциально убывающая с ростом координаты x функция. Считалось, что миграция жертв отсутствует, т. е. $D_N = 0$ — случай, отвечающий часто встречающимся реальным ситуациям, когда подвижность жертв существенно меньше подвижности хищников.

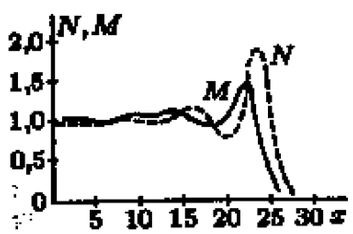


Рис. 112

На «фронте» волны погони формируется пик плотности как хищников, так и жертв (см. рис. 112, где изображены профили функций $M(x,t)$ и $N(x,t)$ в некоторый момент времени). За фронтом волны в системе устанавливается квазиравновесие и значения величин близки к постоянным. Временная динамика образования профиля плотности хищников (рис. 113) свидетельствует о близости процесса на рассматриваемом промежутке времени к автомодельному (самоподобному) режиму. Картина практически без изменений воспроизводится в разные моменты времени в различных областях пространства, причем скорость перемещения характерных точек профиля слабо зависит от времени (амплитуда максимума несколько уменьшается с течением времени, так как волна движется по убывающему «фону» плотности жертв). Такое развитие отвечает автомодельности типа бегущей волны (см. п. 2 § 1 гл. V), в случае которой все величины зависят от комбинации $\xi = x - Dt$, $D > 0$ (для некоторых частных значений параметров автомодельное решение удается получить аналитически).

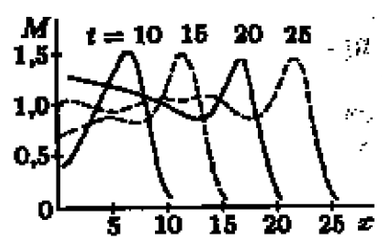


Рис. 113

Изучение распределенных биологических систем хорошо иллюстрирует также и другие соотношения между моделями разных иерархических уровней — точечными и пространственными. Рассмотрим часто используемую модификацию приведенных выше уравнений:

$$\frac{dN}{dt} = \alpha N^2 \frac{N_0 - N}{N_0} - cMN + D_N \frac{\partial^2 N}{\partial x^2},$$

$$\frac{dM}{dt} = (-\beta + \gamma N) M + D_M \frac{\partial^2 M}{\partial x^2},$$

где $N_0 > 0$. Она отличается от классической модели Лотки—Вольтерра видом членов, описывающих динамику жертвы в отсутствие хищника:

1) при малых плотностях популяции (для видов, размножающихся половым путем) скорость роста численности пропорциональна частоте контактов между особями, т. е. квадрату ее плотности (ср. с § 6 гл. I);

2) существует устойчивая равновесная плотность популяции жертв $N = N_0$, определяемая уровнем доступных ресурсов (ср. с логистической моделью в п. 5 § 1 гл. I).

Точечный ($D_N = D_M = 0$) аналог данной модели представляет собой нелинейную систему двух обыкновенных дифференциальных уравнений. Она в соответствующем диапазоне параметров имеет предельный цикл — такую конфигурацию интегральных кривых в фазовой плоскости N, M, когда при $t \to \infty$ траектории «навиваются» на ограниченную замкнутую кривую (автоколебания, процесс качественно подобен описанному в п. 1 § 3 гл. IV).

Получаемое из вычислительного эксперимента, проводимого в том же диапазоне параметров, поведение численности популяций существенным образом зависит от начального состояния системы. Эксперимент отвечал условиям кольцевого ареала (окрестности береговых линий замкнутых водоемов, уровни постоянной высоты в горной местности и т. п.), и поэтому граничные условия для функций $N(x,t)$, $M(x,t)$ брались периодическими по x. Считалось также, что подвижность хищников гораздо больше подвижности жертв ($D_N \ll D_M$).

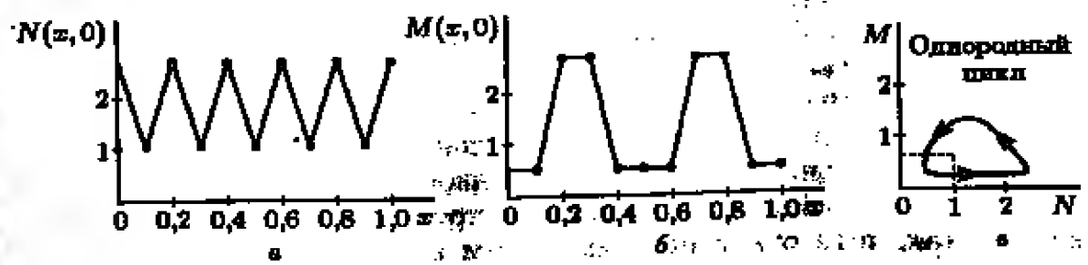


Рис. 114

На рис. 114 в безразмерных единицах показан вид некоторых начальных распределений жертвы (*а*) и хищника (*б*), приводящих к установлению со временем однородного по пространству автоколебательного процесса, соответствующего устойчивому предельному циклу в точечной модели (*в*). Биологическая интерпретация этого результата такова: в однородном ареале («лесу») небольших размеров («зайцы» и

«волки» много раз успевают обежать его за время существования особей) популяции могут взаимодействовать лишь в автоколебательном режиме с изменяющимися в противофазе численностями.

Если же при прочих равных условиях «зайцы» живут оседло, т. е. в одних и тех же местах, а «волки» активно мигрируют по всему лесу в поисках пропитания, то, помимо однородных автоколебаний, возможны пространственно неоднородные (но стационарные по времени) распределения численности популяций.

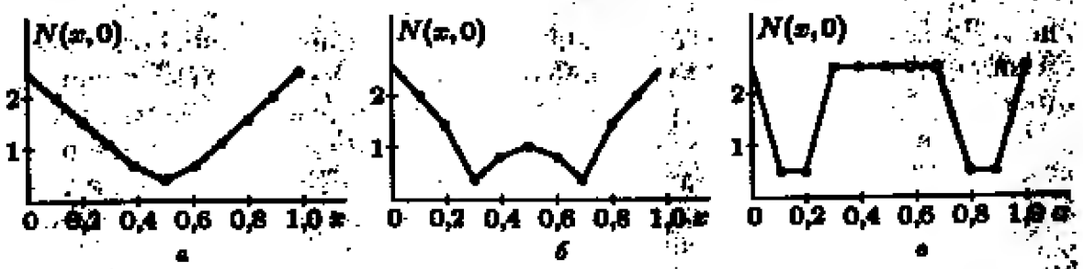


Рис. 115

На рис. 115 *а–в* продемонстрированы некоторые из начальных профилей численности популяции жертв, отвечающих этому не описываемому точечной моделью режиму (рис. 116). В нем численность «волков» одинакова по всему лесу, а «зайцев» на одном конце леса всегда очень много — примерно столько же, сколько их бывает на пике однородных автоколебаний, а на другом конце их всегда очень мало.

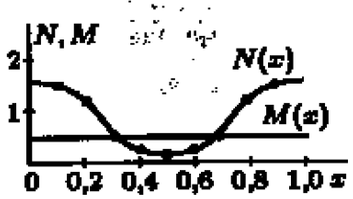


Рис. 116

Несмотря на одинаковые по ареалу условия проживания, в системе появляется так называемая стационарная диссипативная структура (СДС). Заметим, что в отличие от структур горения (§ 3 гл. V), возникновение СДС обусловлено не режимами с обострением и локализацией диффузионных процессов, а конкуренцией источников и стоков энергии (в терминах теории теплопередачи). Неоднородности СДС отвечают реально наблюдаемой «пятнистости» заселения однородных территорий соперничающими биологическими видами.

2. Процессы в переходной экономике. Трансформация экономической системы с доминированием государственной собственности и плановых централизованных начал к конкурентной рыночной экономике с преобладанием частной собственности (либо наоборот) сопровождается сложнейшими переходными процессами. Это неудивительно, так как, хотя оба типа этих систем не существуют в «чистом» виде, разница между ними носит фундаментальный характер: регулирование с помощью финансовых инструментов, реагирующих на изменения экономических показателей, или путем приказов, следующих из анализа возникающих дефицитов; стремление максимизировать прибыль предприятия или выполнить план; почти полное самообеспечение трудоспособных индивидуумов или гарантированный, по мере возможности, минимум государственного социального обеспечения всего населения и т. д.

Математические модели рыночной экономики давно разрабатываются и относительно хорошо изучены, чего нельзя сказать о моделях плановой и, тем более, переходной экономики. Последние не могут быть (даже в принципиальном плане) сведены к моделям классического типа, например к тем, что рассматривались в § 2 гл. IV, поскольку они должны отражать в себе основные черты обеих экономических систем. Эффективный методологический подход к построению моделей, обладающих этим синтетическим свойством, состоит в том, что сначала строятся модели балансов материальных и финансовых потоков, которые в определенном смысле универсальны, т. е. пригодны для описания экономики любого типа. Они «преднамеренно» незамкнуты, а способ их замыкания прямым образом зависит от поведения экономических агентов, политики государства и т. д. При задании разных видов производственных отношений (сценариев) и тем самым при задании разных способов замыкания получаются модели для разных типов экономик.

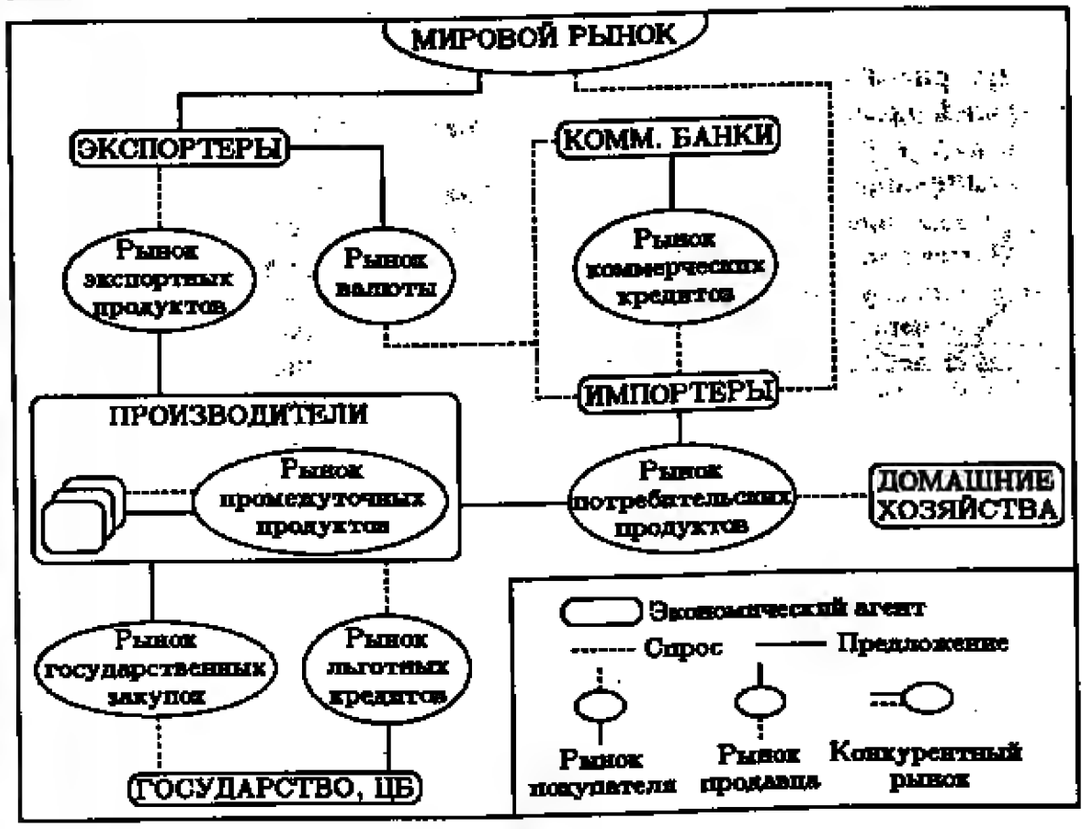


Рис. 117

Блок-схема одной из них приведена на рис. 117. В ней на макроуровне отражены довольно сложные взаимоотношения экономических партнеров, регулирующие производство, обмен и распределение продуктов и услуг, сложившиеся в экономике России к концу первой трети 90-х годов XX века. Видно, что модель отвечает смешанной, переходной экономике: помимо государства (основного агента плановой системы), в ней фигурируют, например, коммерческие банки, работающие в условиях конкуренции с целью извлечения прибыли.

Не описывая полностью все предположения о производственных отношениях, заложенные в модель, охарактеризуем некоторые из них:

1) выделяются секторы, испытывающие конкуренцию импорта, и экспортные отрасли;

2) трудовые коллективы и администрация в секторах заинтересованы в увеличении фонда заработной платы и, несмотря на сокращение спроса на продукцию, добиваются этого с помощью взаимных неплатежей и льготных кредитов Центрального Банка (ЦБ); чистые инвестиции отсутствуют, производственные мощности уменьшаются;

3) изменение условий производства влияет на заработную плату, но не на уровень занятости; банкротств предприятий нет, номинальная безработица невелика;

4) экспортируется только сырье, а импортируются только потребительские товары;

5) рынки контролируются промышленно-финансовой олигархией, вершину которой занимают экспортеры;

6) макроэкономическая политика государства сводится к определению ставок налогов, объемов льготных кредитов ЦБ, государственных закупок, выплат населению из госбюджета и дотаций предприятиям и т. д.

Сформулированный сценарий воплощается в общую модель, в результате чего получается конкретная модель переходного периода. В математическом отношении она представляет собой громоздкую и сложную систему нелинейных обыкновенных дифференциальных уравнений (дополненных большим числом алгебраических уравнений) относительно нескольких десятков основных экономических величин (например, выпусков различных видов продукции) и содержат много определяющих решения характеристик и параметров (например, инфляционные ожидания населения). Эти входные данные находятся и уточняются, как и сценарии, по текущему состоянию системы.

Например, в одном из вариантов модели считалось, что ЦБ не проводит операций на внутреннем валютном рынке, тогда к концу 1993 г. курс доллара должен был бы, согласно модели, подняться до 4000 р./долл. Однако с середины 1993 г. ЦБ начал соответствующие действия, и в реальности курс дошел «лишь» до отметки 1300 р./долл. В модель были внесены учитывающие новую политику изменения, и даваемые ею временные ряды с этого момента неплохо совпадают с фактическими (см. табл. 5).

Вычислительные эксперименты как с этой, так и с другими моделями трансформирующейся экономики, построенными подобным же образом, позволили сделать ряд довольно общих важных выводов. В частности, было установлено, что переход от почти развалившейся в конце 80-х—начале 90-х годов плановой советской экономики к эффективному равновесному состоянию новой рыночной экономической системы даже в лучшем случае займет не менее десяти лет, будет сопровождаться высокой структурной безработицей и банкротствами многих предприятий.

Таблица 5

Дата	Биржевой курс	Расчетный курс	Дата	Биржевой курс	Расчетный курс
03.05	1820	1779	05.07	1998	1984
05	1854	1814	07	2011	2009
10	1859	1809	12	2020	2018
12	1869	1837	14	2022	2029
17	1877	1827	19	2028	2038
19	1881	1858	26	2052	2065
24	1895	1852	28	2052	2076
26	1901	1868	02.08	2060	2087
31	1916	1881	04	2081	2104
02.06	1918	1896	09	2087	2104
07	1940	1891	11	2108	2128
09	1952	1915	16	2117	2137
14	1952	1927	18	2141	2148
16	1959	1927	23	2161	2157
21	1971	1947	25	2156	2171
23	1977	1957	30	2153	2183
28	1985	1969	01.09	2204	2193
30	1989	1983			

Другой, не менее значимый результат экспериментирования с моделями заключается в том, что установлено «попадание» послереформенной российской экономики в особый тип квазиравновесного состояния, отличного от изучаемых в классических политэкономических моделях. Оно весьма неэффективно: в этом состоянии экономическим агентам нет смысла (выгоды) ни сберегать ресурсы, ни инвестировать их в производственный сектор экономики, но зато крайне выгодно сохранять взаимные неплатежи и другие задержки в обращении финансов. Льготные кредиты ЦБ при правильно взвешенной и точно адресованной дозировке несколько улучшают эффективность данного равновесия, но не могут кардинально изменить общую картину (для частных вариантов модели существование такого равновесия устанавливается, как и в § 2 гл. IV, относительно простыми аналитическими методами и описывается несложными установившимися решениями). Оба описанных вывода вполне согласуются с наблюдаемой в последние годы макроэкономической ситуацией в России.

С помощью модели проводятся также и более детальные исследования различных конкретных вопросов текущей экономической политики. К ней предъявляется естественное требование «безопасности», трактуемое в модели как недопустимость резкого разрушения пусть не слишком эффективных, но сложившихся и реально существующих экономических отношений и структур. Это отнюдь не надуманная проблема, поскольку речь идет не о чьем-то сознательном стремлении к разрушению, а о «непрофессиональном» использовании экономических инструментов в очень сложной и неустойчивой ситуации.

Типичная задача — определение размера льготных кредитов, даваемых государством производителям фактически под отрицательный

процент. Модель показала — крайности весьма опасны. Отсутствие льготных кредитов приводит не только к резкому (в течение недель и месяцев) подавлению инфляции (и даже к дефляции), но также и к разрушению производственных структур, в большинстве своем уже приспособившихся к инфляции. Их доходы сокращаются настолько, что неизбежно массовое «бегство» работников предприятий и усиление спада производства. В противоположном случае очень больших льготных кредитов и порождаемой ими гиперинфляции развивается система коммерческих банков. Они планируют свою прибыль, исходя из темпов инфляции. Пока ее рост не слишком велик, их действия, основанные даже на грубых прогнозах, обеспечивают устойчивую прибыль. При гиперинфляции неизбежная неточность прогнозов приводит к систематическим убыткам банков и фактическому «исчезновению» (в относительном смысле) их собственного капитала.

Упомянем также и о двух других конкретных актуальных событиях, значимых для российской экономики и анализировавшихся с помощью вычислительных экспериментов с моделью. Первое из них — «черный вторник» 11 октября 1994 г. — катастрофическое падение курса рубля по отношению к доллару, который через несколько дней возвратился примерно на прежний уровень. Достаточная адекватность модели позволила не только (постфактум) описать динамику основных экономических макропоказателей после «вторника», но и обоснованно определить экономических агентов, которые, пусть невольно, выиграли (базовые отрасли, доходы госбюджета) и проиграли (основная часть населения, импортеры) в результате этого события.

Второе — армейская операция в Чечне, начатая в конце 1994 г. и потребовавшая значительных дополнительных государственных расходов на ее проведение и на мероприятия по восстановлению экономики и социальной сферы республики (по разным оценкам от нескольких триллионов до десятков триллионов рублей). Основной вывод из результатов моделирования: хотя «чеченский кризис» и не может вызвать гиперинфляцию, но, даже при жесткой антиинфляционной политике государства, он вносит заметный вклад в инфляцию и способствует снижению реальных доходов большинства населения.

3. Тоталитарные и анархические эволюции распределения власти в иерархиях. Результаты анализа математической модели системы «власть—общество», полученные в § 4 гл. IV, касались в основном стационарных распределений власти между звеньями иерархии в условиях правового общества. Несмотря на феномен выхода иерархии за рамки предписываемых ей полномочий, реализующийся при рассогласовании характеристик системы, в правовом случае всегда имеет место динамическая устойчивость. Под этим подразумевается следующее: любое нестационарное распределение власти рано или поздно «возвращается» к стационару и тем самым приходит в правовое поле (считается, что сам стационар принадлежит правовой области).

Вычислительный эксперимент наглядно демонстрирует это свойство правовой системы (в некотором смысле оно заложено в ее определении). На рис. 118 показан пример подобного «возврата» власти

в правовое поле для случая базовой модели (п. 4 § 4 гл. IV) с параметрами $\varkappa_0 = 5 \cdot 10^{-3}$, $b = 0{,}9$, $\alpha = 0{,}75$, $k_1 = H = l = 1$, $t_0 = 0$. Начальные данные $p_0(x)$ взяты в виде «ступеньки»: при $0 \leqslant x \leqslant 0{,}3$ имеем $p_0(x) = 2 > p_2$, при $x > 0{,}3$ их значения отвечают стационарному решению. Функция $p(x,t)$ для моментов t_1, t_2, t_3, t_4 изображена сплошными линиями соответственно *1*, *2*, *3*, *4*.

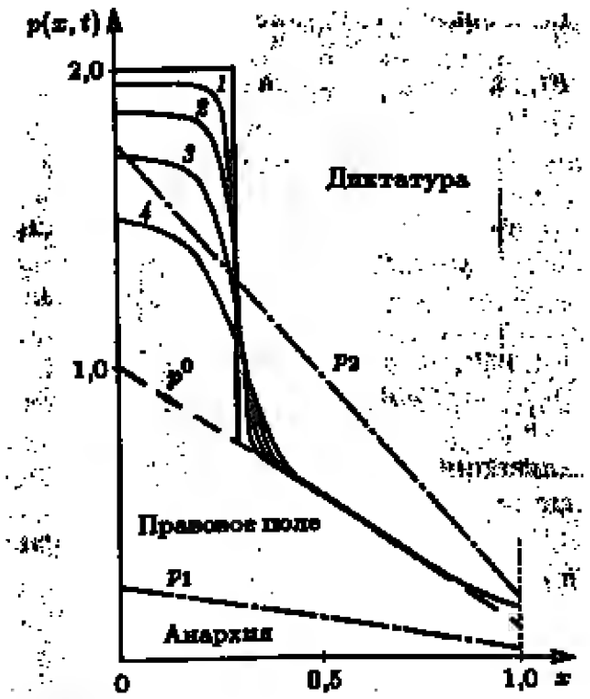


Рис. 118. $t_1 = 0{,}05$, $t_2 = 0{,}15$, $t_3 = 0{,}35$, $t_4 = 0{,}75$

Этот сценарий отвечает тому, что иерархия (в данном случае ее высшие звенья) по каким-то причинам заметно превысила свои максимальные полномочия. Однако реакция общества обеспечила возврат решения в правовые рамки и установление через какое-то время стационарного распределения власти (напомним, что ради простоты здесь считается, что реакция иерархии равна нулю, т. е. «чиновникам» безразличен уровень реализуемой ими власти и они послушно следуют за реакцией общества). Для времени установления с хорошей точностью справедливо $t_{уст} \sim 1/k_1$, т. е. при остальных постоянных параметрах системы распределение власти тем быстрее приходит в правовую область, чем больше интенсивность реакции общества k_1.

Если партнеры в системе «власть—общество» хотят как можно быстрее преодолеть начальную ситуацию, то им следует позаботиться (при прочих равных условиях) об усилении реакции общества (выборы, опросы, информирование и т. п.). Данное заключение вполне соответствует стандартным политологическим рецептам.

Из вычислительных экспериментов (в сочетании с теоретическим анализом), проводимых для общей модели правовой системы, следует ее устойчивость. Другими словами, в правовом случае иерархическая структура и гражданское общество находятся в состоянии устойчивого динамического равновесия.

Совершенно иначе может обстоять дело при некоторых, внешне, быть может, незначительных деформациях общественного сознания по отношению к правовому. В этих случаях нельзя даже приблизительно говорить о существовании какого-либо динамического равновесия системы «власть—общество». Ниже рассмотрен ряд иллюстрирующих это положение сценариев эволюции распределения власти, полученных с помощью вычислительного эксперимента.

I. *Тоталитарная ловушка.* По отношению к правовой системе общественное сознание (реакция общества $F(p, x)$) деформировано так, как изображено на рис. 119 (график $F(p)$ состоит из отрезка линии *1* от $p = 0$ до $p = p_2$ и линии *2*). Это означает, что при $p < p_2$ реакция общества правовая, при $p_2 < p < p_{кр} = 2{,}5 \cdot p^0$ «слаборавовая», т. е. сопротивление превышению власти существует, но уменьшается с ростом p. Наконец, при $p > p_{кр}$ общество «требует» реализации все большей власти и эти требования тем сильнее, чем больше величина p.

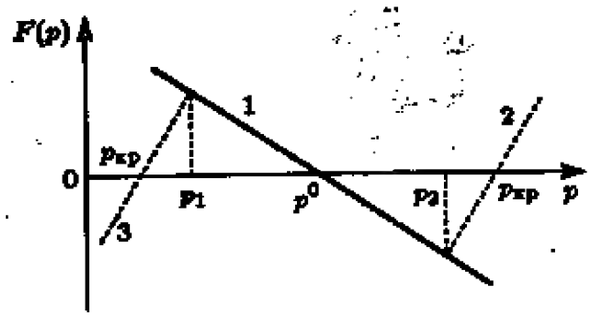


Рис. 119

На рис. 120 приведены результаты расчета по базовой модели (но с $F(p)$ из рис. 119) при $\varkappa_0 = 7{,}5 \cdot 10^{-2}$ (остальные параметры те же, что и на рис. 118). Начальный уровень власти в любом звене иерархии равен нулю ($p_0(x) = 0$, $0 \leqslant x \leqslant 1$). Опишем стадии развития сценария:

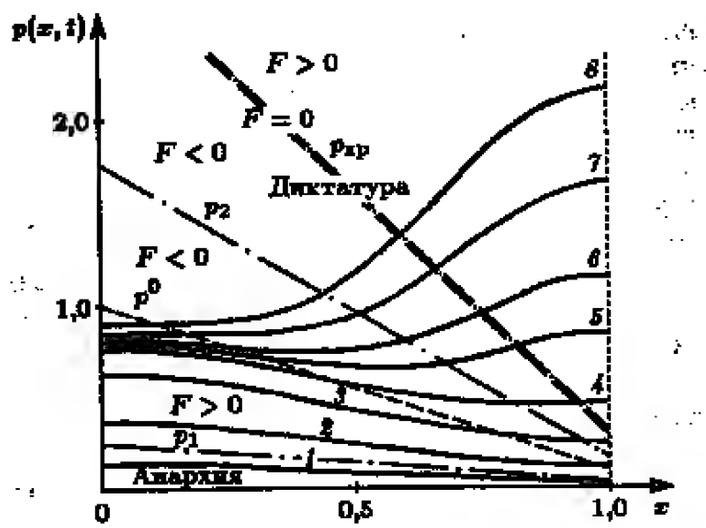


Рис. 120. $t_1 = 0{,}155$, $t_2 = 0{,}635$, $t_3 = 1{,}915$, $t_4 = 6{,}395$, $t_5 = 15{,}995$, $t_6 = 17{,}915$, $t_7 = 19{,}195$, $t_8 = 19{,}835$

а) поскольку при $t_0 = 0$ $p_0(x) \equiv 0$, то в начальный момент времени реакция общества обеспечивает рост величины власти и ее возврат в область $p_1 < p < p_2$ (кривые *1*, *2*), причем большая часть звеньев иерархии находится в области $p_1 < p < p^0$ (кривая *3*) и по-прежнему получает от общества «порции» власти. Решение растет всюду, так как интенсивно работающий механизм передачи власти от старших к младшим звеньям обеспечивает его увеличение в окрестности точки $x = 1$, где реакция общества уже отрицательна (для части кривой *3*, расположенной в области $p^0 < p < p_{кр}$);

б) ситуация, изображенная кривой *4*, ключевая для понимания данного сценария. Благодаря «пологости» профиля власти (параметр $\varkappa_0$ достаточно велик) небольшое количество низших инстанций попадает в область, где $p > p_{кр}$ и где они начинают получать «подпитку» властью в основном от общества ($F(p) > 0$ при $p > p_{кр}$), а не от старших соседей;

в) подобная подпитка приводит (к не очень большому) повышению власти низших инстанций по отношению к уровню власти ближайших старших звеньев (которые пока еще находятся в области $F(p) < 0$, $p^0 < p < p_{кр}$). По постулату (п. 2 § 4 гл. IV) старшие парируют это повышение, «отбирая» часть власти у младших и выходя за счет этого «источника» в область $p > p_{кр}$, $F(p) > 0$ (кривые *5*, *6*);

г) «волна» превышения власти и выхода в область $p > p_{кр}$ распространяется по иерархии справа налево (снизу вверх), и функция $p(x, t)$ необратимо растет до сколь угодно больших значений (кривые *7*, *8* и т. д.) при всех $0 \leqslant x \leqslant 1$ (именно таков смысл употребляемого в данном пункте слова «тоталитарный»).

Описанный сценарий весьма нетривиален (даже с чисто математической точки зрения). Действительно, иерархия, полностью находившаяся в области малых (нулевых) значений $p(x, t)$, с течением времени полностью перешла в область неограниченных значений $p(x, t)$, «преодолев» своеобразную «полосу сопротивления», выставленную законодательством и реакцией гражданского общества (область $p^0 < p < p_{кр}$). Сработали два фактора — деформация общественного сознания и чересчур интенсивные механизмы перераспределения власти в иерархии. Именно последние «вытянули» всю иерархию звено за звеном в область $p > p_{кр}$, куда первоначально попала лишь незначительная часть низших инстанций (ловушка). При уменьшении величины $\varkappa_0$ ситуация нормализуется — решение выходит на стационар.

II. *Анархическая ловушка.* В предыдущем сценарии (как и в модели вообще) не заложено никаких сознательных поползновений иерархической структуры на «диктатуру». При обратной деформации общественного сознания реализуется обратный сценарий. Функция $F(p)$ изображена на рис. 119 частью линии *1* при $p > p_1$ и линией *3*. При $p > p_1 = 0{,}85\,p^0$ реакция как и в правовом случае, при $p_{кр} = 0{,}7\,p^0 < p < p_1$ реакция «слаборавовая», т. е. положительна, но уменьшается с уменьшением p. Наконец, при $p < p_{кр}$ реакция общества направлена на уменьшение власти и тем больше, чем больше это уменьшение.

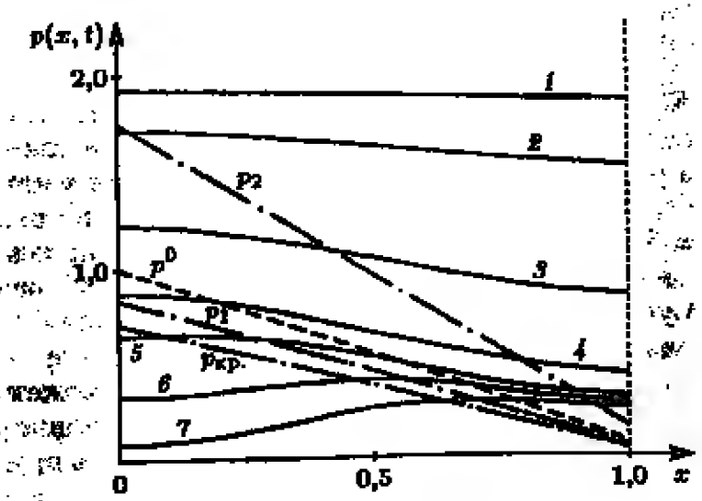


Рис. 121. $t_1 = 0{,}075$, $t_2 = 0{,}315$, $t_3 = 1{,}275$, $t_4 = 3{,}195$, $t_5 = 8{,}315$, $t_6 = 10{,}875$, $t_7 = 11{,}515$

Результаты расчета приведены на рис. 121. Параметры модели те же, что и в случае рис. 120 (за исключением, естественно, вида $F(p)$). Начальные данные: $p_0(x) = 2$, $0 \leqslant x \leqslant 1$, т. е. в момент $t_0 = 0$ распределение власти полностью находится в области $p > p^0$ (где $F(p) < 0$).

Эволюция распределения власти идет по обратному по сравнению со случаем I сценарию: решение при достаточно больших t обращается в нуль для всех $0 \leqslant x \leqslant 1$, проходя полосу $p_{кр} < p < p^0$, где $F(p) > 0$. Вначале уровень власти уменьшается, и большая часть профиля власти попадает в правовую область $p_1 < p < p_2$. Затем часть старших инстанций незначительно «проваливается» в область $p < p_{кр}$ (линия *5*), однако этого достаточно для того, чтобы со временем в эту область перешли все звенья иерархии и произошло необратимое уменьшение решения $p(x,t)$ до нуля при всех $0 \leqslant x \leqslant 1$; как и в случае I, механизмы перераспределения власти слишком сильны, их уменьшение приводит к нормальной эволюции к стационарному решению.

III. *Усталость общества и «дремлющая» ловушка.* При заданных в пп. I, II параметрах системы «власть—общество» эволюция распределения власти предопределена в том смысле, что для любых начальных данных $p_0(x)$ реализуются описанные выше сценарии (за исключением количественных различий качественное поведение функции $p(x,t)$ то же самое).

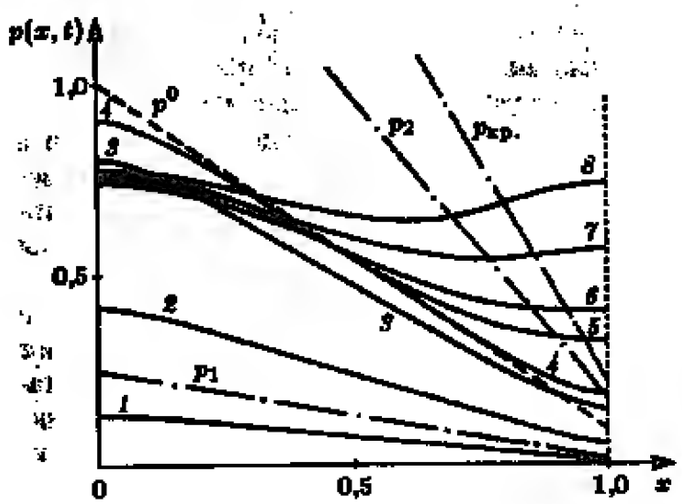


Рис. 122. $t_1 = 0{,}155$, $t_2 = 0{,}63$, $t_3 = 2{,}555$, $t_4 = 7{,}035$, $t_5 = 21{,}115$, $t_6 = 55{,}675$, $t_7 = 99{,}835$

Более сложная и разнообразная эволюция может реализовываться в случаях, когда характеристики системы изменяются со временем.

На рис. 122 приведены результаты расчета модели с функцией $F(p)$ такой же, как и для расчета на рис. 120, но с $\varkappa_0 = 10^{-2}$ (в этом случае стационарное решение для правовой системы слегка выходит в область $p > p_2$, но так, что между ним и областью $p > p_{кр}$ остается заметный «зазор»). Второе отличие в том, что в момент $t = t_{кр} = 8$ как положительная, так и отрицательная части реакции общества по амплитуде скачком уменьшаются в 10 раз, хотя качественно, как функции p, они имеют прежний вид (линия *1* при $p < p_2$ и линия *2* на рис. 119).

Такое изменение $F(p)$ со временем может быть интерпретировано как усталость общества. Эволюция функции $p(x,t)$ следующая:

а) решение растет и входит в область $p_1 < p < p_2$, как и на рис. 120 (кривые *1*, *2*);

б) решение на некоторое время становится близким к стационарному (кривые *3*, *4*), которое, в отличие от сценария I, при данном значении $\varkappa_0$ существует;

в) уменьшение реакции общества в момент $t = t_{кр}$ приводит к пространственному «разглаживанию» распределения власти и попаданию части низших инстанций в область $p > p_{кр}$ — кривая *5* (решение оставалось бы стационарным при прежней $F(p)$);

г) ловушка, связанная с видом $F(p,x,t)$, «просыпается», и далее решение неограниченно растет при всех $0 \leqslant x \leqslant 1$ (кривые *6* и т. д.).

IV. *Активизация общества и «возрождающийся» стационар.* В отличие от предыдущего расчета, в моменты времени $6 < t < 10$ амплитуда реакции возрастает в 10 раз (причем величина $p_{кр}$ также увеличивается). Такое поведение $F(p)$ может быть интерпретировано как активизация общества.

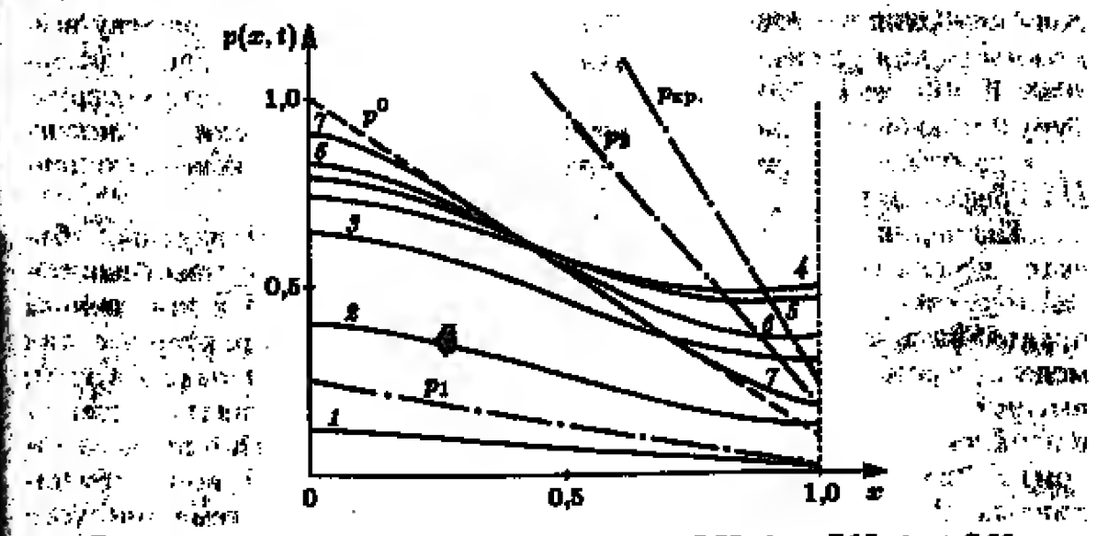


Рис. 123. $t_1 = 0{,}15$, $t_2 = 0{,}75$, $t_3 = 2{,}35$, $t_4 = 5{,}55$, $t_5 = 7{,}15$, $t_6 = 7{,}55$, $t_7 = 19{,}90$

Решение (рис. 123) вначале ведет себя как и в случае рис. 122, попадая в область $p > p_{кр}$ (кривые *3*, *4*). При взятом значении $\varkappa_0 = 10^{-1}$ и не меняющейся со временем функции $F(p)$ функция $p(x,t)$ неограниченно росла бы (как и на рис. 122). Однако увеличение амплитуды $F(p)$ при $t > 6$ приводит к тому, что в системе появляется стационарное решение, на которое с течением времени (кривые *5*, *6*) распределение власти и «выходит» (кривая *7*), избежав грозившей ему ловушки.

Разумеется, сценарии I–IV и их политологические интерпретации весьма условны (хотя бы потому, что при заметном превышении или принижении власти модель, строго говоря, неприменима, так как нарушается предположение о законопослушности системы).

Однако они демонстрируют потенциальное богатство возможных поведений распределения власти в иерархии, заложенных в модель, которая содержит различные прямые и обратные связи, нелинейности и пространственно-временные вариации характеристик объекта, а также указывают на возможность осмысленного толкования эволюции изучаемой системы.

Например, анархическая ловушка может быть соотнесена с российскими событиями 1917 г. (в большей степени) и конца 80-х—начала 90-х годов (в меньшей степени). Действительно, «провал» в уровне власти высших иерархов (кульминация — отречение монарха и отставка

генсека) под сильнейшим давлением гражданского общества и при «безответственных» действиях иерархии с последующей цепной реакцией уменьшения и даже «исчезновения» властного влияния в остальных ее звеньях вполне описывается этим сценарием. Интерпретация имеет смысл, конечно, лишь на начальной стадии процесса (общая тенденция к «анархии»). Далее она не может быть продолжена хотя бы потому, что «старая» властная структура просто перестала существовать.

Еще один пример — политическое развитие России в последней трети 1993 г. Ему отвечает сценарий «возврат власти в рамки полномочий». Формально выпадшая из (не совсем четких) границ правового поля иерархия в то же время обеспечила гражданскому обществу возможность для усиленного выражения своей реакции с помощью выборов в Парламент. Это позволило обществу «вернуть» властную структуру в правовую область, уточненные рамки которой были установлены к тому же при непосредственном его участии путем конституционного референдума.

Подчеркнем: математические модели трудноформализуемых объектов всегда содержат в себе плохо или не полностью известные поведенческие характеристики живых существ. Поэтому к ним нельзя предъявлять требования адекватности и точности, характерные для моделирования проблем технологии и естествознания. Однако и в этой, высшей степени сложности сфере анализ, прогноз и принятие решений в немалой мере основаны на использовании (пусть почти и не осознанном) каких-то моделей и методов моделирования, чаще всего примитивных. Поэтому широкое применение методологии математического моделирования и вычислительного эксперимента представляется неизбежным и в данной области человеческой деятельности.

Библиография к главе VI: [4, 10, 13, 14, 19, 23, 27, 28, 32, 35, 36, 39, 41, 48, 49, 52–54, 58, 66, 81, 82].

СПИСОК ЛИТЕРАТУРЫ

1. *Абрамов А.П., Иванилов Ю.П.* Физика и математическая экономика. — М.: Знание, 1991. — 32 с.
2. *Адъютов М.М., Клоков Ю.А., Михайлов А.П.* Автомодельные тепловые структуры с сокращающейся полупрямой // Дифференц. уравнения. — 1983. — Т. 19, № 7. — С. 1107–1114.
3. *Айзерман М.А.* Классическая механика. — М.: Наука, 1980. — 368 с.
4. *Александров В.В., Арчипов П.П., Пархоменко В.П., Стенчиков Г.Л.* Глобальная модель системы океан—атмосфера и исследование ее чувствительности к изменению концентрации CO_2 // Известия АН СССР. Физика атмосферы и океана. — 1983. — Т. 19, № 5. — С. 451–458.
5. *Амелькин В.В.* Дифференциальные уравнения в приложениях. — М.: Наука, 1987. — 160 с.
6. *Ануфриева И.А., Михайлов А.П.* Неограниченные решения квазилинейного уравнения переноса. — Препринт / ИПМ АН СССР. — М., 1985. — № 34. — 29 с.
7. *Арсеньев А.А., Самарский А.А.* Что такое математическая физика. — М.: Знание, 1983. — 64 с.
8. *Ахромеева Т.С., Курдюмов С.П., Малинецкий Г.Г., Самарский А.А.* Нестационарные структуры и диффузионный хаос. — М.: Наука, 1992. — 542 с.
9. *Ашманов С.А.* Введение в математическую экономику. — М.: Наука, 1984. — 296 с.
10. *Базыкин А.Д.* Математическая биофизика взаимодействующих популяций. — М.: Наука, 1985. — 182 с.
11. *Баранов В.Б., Краснобаев Н.В.* Гидродинамическая теория космической плазмы. — М.: Наука, 1977. — 336 с.
12. *Баренблатт Г.И.* Подобие, автомодельность, промежуточные асимптотики. — Л.: Гидрометиздат, 1982. — 208 с.
13. *Белоцерковский О.М.* Численное моделирование в механике сплошных сред. — М.: Наука, 1994. — 442 с.
14. *Бочков М.В., Лобачев Л.А., Четверушкин Б.Н.* Химическая кинетика образования NO при горении метана в воздухе // Математическое моделирование. — 1992. — Т. 4, № 9. — С. 3–36.
15. *Будак Б.М., Самарский А.А., Тихонов А.Н.* Сборник задач по математической физике. — М.: Наука, 1980. — 686 с.
16. *Вабищевич П.Н.* Численное моделирование. — М.: Изд-во МГУ, 1993. — 152 с.
17. *Волосевич П.П., Дарьин Н.А., Леванов Е.И., Скиртладзе Н.М.* Задача о поршне в газе с источниками и стоками (автомодельные решения). — Тбилиси: Изд-во ТГУ, 1986. — 238 с.
18. *Волосевич П.П., Леванов Е.И.* Автомодельные решения задач газовой динамики с учетом теплопроводности. — М.: Изд-во МФТИ, 1996. — 212 с.
19. *Гамалий Е.Г., Демченко Н.Н., Лебо И.Г. и др.* Теоретическое исследование устойчивости сжатия тонкостенных оболочечных мишеней, облучаемых лазерами с энергией в импульсе порядка 1 кДж // Квантовая электроника. — 1988. — Т. 15, № 8. — С. 1622–1632.
20. *Гантмахер Ф.Р.* Лекции по аналитической механике. — М.: Наука, 1968. — 300 с.
21. *Годунов С.К., Рябенький В.С.* Разностные схемы (введение в теорию). — М.: Наука, 1977. — 440 с.

22. *Головизнин В.М., Самарский А.А., Фаворский А.П.* Вариационный подход к построению конечноразностных математических моделей в гидродинамике // ДАН СССР. — 1977. — Т. 235, № 6. — С. 1285–1288.
23. *Гольдин В.Я., Анистратов Д.Ю.* Реактор на быстрых нейтронах в саморегулируемом нейтронно-ядерном режиме // Математическое моделирование. — 1995. — Т. 7, № 10. — С. 21–32.
24. *Демидов М.А., Михайлов А.П.* Эффекты локализации и образования структур при сжатии конечной массы газа в режиме с обострением // ПММ. — 1986. — Т. 50, № 1. — С. 119–127.
25. *Дородницын А.А.* Информатика: предмет и задачи // Кибернетика. Становление информатики. — М.: Наука, 1996.
26. *Дородницын В.А., Елеин Г.Г.* Симметрия в решении уравнений математической физики. —М.: Знание, 1984. — 64 с.
27. *Елизарова Т.Г., Павлов А.Н., Четверушкин Б.Н.* Применение кинетического алгоритма для расчета газодинамических течений // Дифференц. уравнения. — 1985. — Т. 21, № 7. — С. 1179–1185.
28. *Елизарова Т.Г., Павлов А.Н., Четверушкин Б.Н.* Использование квазигидродинамической системы для расчета обтекания тела с иглой // ДАН СССР. — 1987. — Т. 297, № 2. — С. 327–331.
29. *Елизарова Т.Г., Четверушкин Б.Н.* Кинетические алгоритмы для расчета гидродинамических течений // ЖВМ и МФ. — 1985. — Т. 25, № 10. — С. 1526–1533.
30. *Зельдович Я.Б., Райзер Ю.П.* Физика ударных волн и высокотемпературных гидродинамических явлений. — М.: Наука, 1966. — 688 с.
31. *Змитренко Н.В., Курдюмов С.П., Михайлов А.П.* Теория режимов с обострением в сжимаемых средах // Итоги науки и техники. Современные проблемы математики. Новейшие достижения. Т. 28. — М.: ВИНИТИ, 1986. — С. 3–94.
32. *Змитренко Н.В., Курдюмов С.П., Михайлов А.П., Самарский А.А.* Локализация термоядерного горения в плазме с электронной теплопроводностью // Письма в ЖЭТФ. — 1977. — Т. 26, № 9. — С. 620–624.
33. *Змитренко Н.В., Михайлов А.П.* Инерция тепла. — М.: Знание, 1982. — 64 с.
34. *Ибрагимов Н.Х.* Группы преобразований в математической физике. — М.: Наука, 1983. — 280 с.
35. *Иванова Т.С., Рузмайкин А.А.* Метод решения задачи магнитогидродинамического динамо Солнца // ЖВМ и МФ. — 1976. — Т. 16, № 4. — С. 956–968.
36. *Иванова Т.С., Рузмайкин А.А.* Нелинейная магнитогидродинамическая модель динамо Солнца // Астроном. журн. — 1977. — Т. 54. — С. 846–858.
37. *Калиткин Н.Н., Михайлов А.П.* Идеальное решение задачи зачета взаимных долгов // Математическое моделирование. — 1995. — Т. 7, № 6. — С. 111–117.
38. *Клоков Ю.А., Михайлов А.П.* Об одной краевой задаче Неймана для интегро-дифференциального уравнения // Дифференц. уравнения. — 1996. — Т. 32, № 8. — С. 1110–1113.
39. *Коробейников В.П.* Математическое моделирование катастрофических явлений природы. — М.: Знание, 1986. — 48 с.
40. *Краснощеков П.С., Петров А.А.* Принципы построения моделей. — М.: Изд-во МГУ, 1983. — 264 с.
41. *Крыскин Ю.А., Самарская Е.А., Тишкин В.Ф.* Балансовая модель распространения примеси в плановом фильтрационном потоке // Математическое моделирование. — 1993. — Т. 5, № 6. — С. 69–84.
42. *Ландау Л.Д., Лифшиц Е.М.* Механика. — М.: Наука, 1973. — 207 с.
43. *Ландау Л.Д., Лифшиц Е.М.* Механика сплошных сред. — М.: Гостехиздат, 1953. — 788 с.
44. *Лейбензон Л.С.* Движение природных жидкостей и газов в пористой среде. — М.–Л.: Гостехиздат, 1947. — 244 с.
45. *Лойцянский Л.Г.* Механика жидкости и газа. — М.: Гостехиздат, 1950. — 676 с.
46. *Марчук Г.И.* Методы вычислительной математики. — М.: Наука, 1989. — 608 с.
47. Математическое моделирование / Под ред. Дж. Эндрюса, Р. Мак-Лоуна; пер. с англ. — М.: Мир, 1979. — 278 с.

48. *Михайлов А.П.* Математическое моделирование распределения власти в иерархических структурах // Математическое моделирование. — 1994. — Т. 6, № 6. — С. 108–138.
49. *Михайлов А.П.* Моделирование эволюции распределения власти в государственных иерархиях // Вестник Фонда «Российский общественно-политический центр». — 1996. — № 2. — С. 26–39.
50. *Моисеев Н.Н.* Математические задачи системного анализа. — М.: Наука, 1981. — 488 с.
51. *Овсянников Л.В.* Групповой анализ дифференциальных уравнений. — М.: Наука, 1978. — 400 с.
52. *Пархоменко В.П., Стенчиков Г.Л.* Математическое моделирование климата. — М.: Знание, 1986. — 32 с.
53. *Петров А.А.* Экономика. Модели. Вычислительный эксперимент. — М.: Наука, 1996.
54. *Петров А.А., Поспелов И.Г., Шананин А.А.* Опыт математического моделирования экономики. — М.: Энергоиздат, 1996. — 544 с.
55. *Полубаринова-Кочина П.Я.* Теория движения грунтовых вод. — М.: Наука, 1977.
56. *Понтрягин Л.С., Болтянский В.Г., Гамкрелидзе Р.В., Мищенко Е.Ф.* Математическая теория оптимальных процессов. — М.: Гостехиздат, 1961. — 392 с.
57. *Попов Ю.П., Самарский А.А.* Вычислительный эксперимент. — М.: Знание, 1983. — 64 с.
58. *Рязиченко Г.Ю., Рубин А.Б.* Математические модели биологических репродукционных процессов. — М.: Изд-во МГУ, 1993. — 300 с.
59. *Самарский А.А.* Введение в численные методы. — М.: Наука, 1982. — 272 с.
60. *Самарский А.А.* Математическое моделирование и вычислительный эксперимент // Вестник АН СССР. — 1979. — № 5. — С. 38–49.
61. *Самарский А.А.* Теория разностных схем. — М.: Наука, 1989. — 616 с.
62. *Самарский А.А., Андреев В.Б.* Разностные методы для эллиптических уравнений. — М.: Наука, 1976. — 352 с.
63. *Самарский А.А., Галактионов В.А., Курдюмов С.П., Михайлов А.П.* Режимы с обострением в задачах для квазилинейных параболических уравнений. — М.: Наука, 1987. — 478 с.
64. *Самарский А.А., Гулин А.В.* Устойчивость разностных схем. — М.: Наука, 1973. — 416 с.
65. *Самарский А.А., Колдоба А.В., Повещенко Ю.А. и др.* Разностные схемы на нерегулярных сетках. — Минск: ЗАО «Критерий», 1996. — 274 с.
66. *Самарский А.А., Михайлов А.П.* Компьютеры и жизнь (математическое моделирование). — М.: Педагогика, 1987. — 128 с.
67. *Самарский А.А., Николаев Е.С.* Методы решения сеточных уравнений. — М.: Наука, 1978. — 592 с.
68. *Самарский А.А., Попов Ю.П.* Разностные схемы газовой динамики. — М.: Наука, 1980. — 352 с.
69. *Седов Л.И.* Механика сплошной среды. Т. I. — М.: Наука, 1973. — 536 с.
70. *Седов Л.И.* Методы подобия и размерности в механике. — М.: Наука, 1981. — 448 с.
71. *Сидоров А.Ф., Шанесов В.П., Яненко Н.Н.* Метод дифференциальных связей и его приложения к газовой динамике. — Новосибирск: Наука, 1984. — 272 с.
72. *Сычев В.П.* Введение в кинетическую теорию газов. — М.: Наука, 1971. — 332 с.
73. *Тихонов А.Н., Костомаров Д.П.* Вводные лекции по прикладной математике. — М.: Наука, 1984. — 190 с.
74. *Тихонов А.Н., Самарский А.А.* Уравнения математической физики. — М.: Наука, 1972. — 736 с.
75. *Федоренко Р.П.* Введение в вычислительную физику. — М.: Изд-во МФТИ, 1994. — 526 с.
76. *Aula X.J.R.* Mathematical Modeling // Encyclopedia of Physical Science. — 1987. — V. 7. — P. 719–728.
77. *Bender E.A.* An Introduction to Mathematical Modelling. — N.Y.: Wiley, 1978.
78. *Cross M., Moscardini A.O.* Learning the Art of Mathematical Modelling. — N.Y.: Wiley, 1985. — 154 p.

79. *Dym C.L., Ivey E.S.* Principles of Mathematical Modelling. — N.Y.: Academic Press, 1980. — 256 p.
80. *Hoerner S., von.* Population Explosion and Interstellar Expansion // Journal of British Interplanetary Society. — 1975. — V. 28, No. 11. — P. 691–712.
81. *Jacoby S.L.S, Kowalik J.S.* Mathematical Modelling with Computers. — Englewood Cliffs, N.J.: Prentice Hall, Inc., 1980. — 292 p.
82. *Leko I.G., Rozanov V.B., Tiskkin V.F. et al.* Numerical Simulation of Ruchtmyer—Meshkov Instalility / Eds. R. Young, J. Glimm, B. Boston // Proc. of the Fifth Int. Workshop on Compressible Turbulent Mixing. — N.Y.: Word Scientific, 1995. — P. 346–356.
83. *Lekman R.S.* Computer, Simulation and Modelling: An Introduction. — N.Y.: Wiley, 1977.
84. Mathematical Modelling /Eds. J.G.Andrews, R.R.McLone. — London: Butterworths, 1976.
85. *Rapoport A.* Mathematical Models in the Social and Behavioral Sciences. — N.Y.: Wiley, 1983.
86. *Saaty T.L., Alexander J.M.* Thinking with Models: Mathematical Models in the Physical, Biological and Social Sciences. — N.Y.: Pergamon Press, 1981.
87. *Samarskii A.A., Galaktionov V.A., Kurdyumov S.P., Mikhailov A.P.* Blow-up in Quasilinear Parabolic Equations. — Berlin: Walter de Gruyter, 1995. — 534 p.
88. *Samarskii A.A., Nikolaev E.S.* Numerical Methods for Grid Equations. V. 1, 2. — Basel: Birkhäuser Verlag, 1981. — 242 p., 502 p.
89. *Samarskii A.A., Vabishchevich P.N.* Computational Heat Transfer. V. 1, 2. — N.Y.: Wiley, 1995. — 406 p., 422 p.

A. A. SAMARSKII, A. P. MIKHAILOV

MATHEMATICAL MODELLING: IDEAS. METHODS. EXAMPLES

Physical and Mathematical Literature Publishing Company
Russian Academy of Sciences

Moscow, 1997, 320 pages

ABOUT AUTHORS

Alexander A. Samarskii:
Academician of Russian Academy of Sciences (RAS), Director of Institute of Mathematical Modelling RAS (IMM), Professor of M. V. Lomonosov Moscow State University (MSU), President of IMACS — RUSSIA (the Russian affiliate group of IMACS — International Association for Mathematics and Computers in Simulation).

Was born in 1919 in Donetsk region of Ukraina, graduated from department of Physics of MSU. Academician Samarskii received a Ph. D. in mathematics from the M. V. Lomonosov State University in Moscow in 1948, and a D. Sc. from the M. V. Keldysh Institute of Applied Mathematics of the USSR Academy of Sciences in 1957 (these were for work and publications in the fields of mathematical physics and numerical analysis — specifically difference schemes).

He became an Associate Member of the USSR Academy of Sciences in 1966, and was promoted to Full Member in 1976. He is at present director of IMM RAS, which was founded in 1986 due to his initiative. His scientific and educative career began in the 40th at MSU. A. A. Samarskii served as lecturer, associate professor, professor and, since 1983, he has been head of a chair on the Faculty of Computational Mathematics and Cybernetics. He is also head of a department at the M. V. Keldysh Institute of Applied Mathematics of the USSR Academy of Sciences, a function which he has held from its foundation in 1953 and head of a chair at Moscow Institute of Physics and Technology (MFTI).

Academician Samarskii has been since 1985 a leader of the National Program for Mathematical Modelling in Science and Technology. His research interests includes difference methods, mathematical physics, the theory of nonlinear partial differential equations, plasma physics, mathematical modelling of complex nonlinear systems and phenomena, and computer simulation techniques. He has obtained in those fields outstanding results appreciated by specialists all over the world, and contributed greatly to the development of those subjects.

He has published more then 20 monographs, textbooks and popular science books, including the well-known textbook *Equations of Mathematical Physics* written in 1951 in collaboration with A. N. Tikhonov. This book was reprinted many times in Russia and other countries, it was translated into thirteen languages. His textbook *Theory of Difference Schemes* (in Russian) has been reprinted many times. He is also the author of the monograph written in collaboration with E. S. Nikolaev and published in English under the title *Numerical Methods for Grid Equations* (Birkhäuser, 1989); the monograph *Blow-up in Quasilinear Parabolic Equations* written in collaboration with V. A. Galaktionov, S. P. Kurdyumov and A. P. Mikhailov, and published in English (Walter de Gruyter, 1995); the monograph *Computational Heat Transfer* written in collaboration with P. N. Vabishevich and published in English (Wiley, 1995) and other monographs. He has published more then 400 scientific and popular science papers.

A. A. Samarskii participates actively in Russian and International Journals. He is the Editor-in-Chief of the Journal "Mathematical Modelling" (Russia), and an Associate Editor of the Journals "Differential Equations", "Journal of Computational Mathematics and Mathematical Physics", International Journals "Surveys on Mathematics for Industry" and "Mathematical Models and Methods in Applied Sciences".

His long-time career as an educator has resulted in the creation of a world-known scientific school in the field of Computational Mathematics and Mathematical Modelling

including not only Russian researchers, but also colleagues from Germany, Bulgaria, Hungary and some other countries. There are many highly experienced specialists, professors, associate members and members of the RAS among his students. Academician Samarskii has organized and chaired many international conferences on mathematical physics and difference methods, and served as guest lecturer at many other international forums.

Alexander P. Mikhailov:
Was born in 1947 in Moscow, graduated from MFTI. Head of department at IMM RAS. Professor of M. V. Lomonosov MSU. Doctor of physical and mathematical sciences. Secretary of IMACS — RUSSIA.

Professor Mikhailov received a Ph. D. in mathematics from the M. V. Keldysh Institute of Applied Mathematics (IAM) of RAS in 1978, and a D. Sc. from the M. V. Lomonosov MSU in 1987. Since 1974 till 1990 he was a researcher at IAM RAS, since 1990 he is a head of department at IMM RAS.

His research activity is devoted to mathematical physics, the theory of nonlinear phenomena, mathematical modelling of dissipative structures, problems of mathematical modelling of some socio-political processes. He is the author of more than 100 scientific and popular papers, including book for school-students *Computers and Life* (*Mathematical Modelling*) which was written in collaboration with academician A. A. Samarskii and was published in amount of 200 thousands copies (Pedagogika, Moscow, 1987, in Russian). He is a member of Editorial Board of Russian Journal "Mathematical Modelling", participated as a speaker in many international conferences.

ABSTRACT

It is impossible to imagine contemporary science, technology and many other fields of creative activity apart from mathematical modelling. The essense of this methodology is changing of a real object by its mathematical "image" and studing of this image with the help of computational methods by computers. As a result of computational experiments with mathematical models a researcher obtains a detailed and reliable information concerning original object (phenomenon, process, etc.). Due to its universality, flexibility and accuracy this methodology won a lot of fields of knowledges from technical systems design and control to analysis of social processes.

No wonder that many brilliant books are devoted to various aspects of mathematical modelling. Hence it is necessary to answer following questions: what is the place of this book among others, does it have its own face? There exist important reasons for positive response. As far as it is known, a majority of books on mathematical modeling follows one more or less standard scheme. This scheme includes a general speculations about nature of mathematical modelling and consideration of computer experiments in connection with these or those concrete problems. The later as a rule are determined by professional interests of authors. Such an approach has one evident disadvantage — loss of huge number of potential readers, which are not specialists in problems under consideration.

Authors of this book followed other way. Their aim was to select approaches common for models constructing independently from branch of knowledges. Indeed, a limited number of these approaches exist. The world surrounding us is united. Researches use effectively this gift of the world when they deal with its mathematical description. Of course, contents of the book reflects in some extent author's personal experience. Nevertheless the approach mentioned above helps to extend frameworks of consideration and to represent wide possibilities of mathematical modelling from mechanics to social sciences.

Besides models building authors also paid sufficient attention to their realisation — i. e. to basic principles of construction of model's discrete analogs. In this sense the book is devoted to links between parts *model* and *algorithm* from triad of mathematical modelling *model* — *algorithm* — *code*.

As for genre and style of the book one may determine it as a book for reading rather than textbook or monography, because authors prefered to represent ideas and examples rather than rigorous mathematical procedures, which may be found in specialized books. It should be also noted, that a manner of representation of material corresponds to typical understanding of mathematical modelling existing among Russian professionals, whose advances are world-known.

Authors whould hope that emphasized features of this book will attract attention of specialists on mathematical modelling from various fields as well as interest of teachers and students, that are going to deal with this fascinating creative activity.

AUTHORS' ADDRESSES

A. A. Samarskii: Institute of Mathematical Modelling Russian Academy of Sciences, Miusskaya sq., 4A, 125047, Moscow, Russia.
E-mail — samarskii@imamod.msk.su.
Fax number — (7–095) 972–07–23.

A. P. Mikhailov: Institute of Mathematical Modelling Russian Academy of Sciences, Miusskaya sq., 4A, 125047, Moscow, Russia.
E-mail — mikhailov@imamod.msk.su.
Fax number — (7–095) 972–07–23.

CONTENTS

INTRODUCTION

CHAPTER I
SIMPLIEST MODELS
AND BASIC CONCEPTS OF MATHEMATICAL MODELLING

§ 1. Elementary Mathematical Models
1. Fundamental laws of nature. 2. Variational Principles. 3. Application of Analogies for Models Building. 4. Hierarchical Approach to Models Construction. 5. On Nonlinearity of Mathematical Models. 6. Some Preliminary Conclusions. Exercises.

§ 2. Examples of Models Deduced from Fundamental Laws of Nature
1. Trajectory of Submarine Rising. 2. Deviation of a Charged Particle in Electron Tube. 3. Oscillations of Saturn Rings. 4. Motion of a Ball Connected with Spring. 5. Conclusion. Exercises.

§ 3. Variational Principles and Mathematical Models
1. General Scheme of Hamilton Principle. 2. Third Way for Deducing of System "Ball—Spring" Model. 3. Oscillations of Pendulum in Gravitational Field. 4. Conclusion. Exercises.

§ 4. Example of Hierarchy of Models
1. Various Ways of Given External Force Action to System Ball—Spring. 2. Motion of Bracing Point and a Spring on Rotating Rod. 3. Accounting of Friction. 4. Two Types of Nonlinear Models of Ball—Spring System. 5. Conclusion. Exercises.

§ 5. Universality of Mathematical Models
1. Liquid in U-shape Vessel. 2. Oscillatory Electrical Contour. 3. Small Oscillations at Interference of Two Biological Populations. 4. Simpliest Model of Changing of Wages and Employment. 5. Conclusion. Exercises.

§ 6. Some Models of Simpliest Nonlinear Objects
1. About Origin of Nonlinearity. 2. Three Modes in Nonlinear Model of Population. 3. An Influence of Strong Nonlinearity to Process of Oscillations. 4. On numerical Methods. Exercises.

CHAPTER II
OBTAINING OF MODELS
FROM FUNDAMENTAL LAWS OF NATURE

§ 1. Conservation of Matter
1. Stream of Particles in Tube. 2. Main Assumptions on Gravitational Flow of Underground Waters. 3. Matter Balance in Element of Ground. 4. Law of Conservation of Matter Closure. 5. On Some Properties of Boussinesq Equation. Exercises.

§ 2. Conservation of Energy
1. Preliminary Informations about Heat Transfer Processes. 2. Deduction of Fourier Law from Molecular-Kinetic Considerations. 3. Equation of Heat Balance. 4. Posing of Typical Boundary Conditions for Equation of Heat Conductivity. 5. On Peculiarities of Heat Transfer Models. Exercises.

§ 3. Conservation of Number of Particles
1. Main Concepts of Heat Radiation Theory. 2. Equation of Photons Balance in Medium. 3. Some Properties of Radiation Transfer Equation. Exercises.

§ 4. Combined Application of Several Fundamental Laws
1. Preliminary Concepts of Gasdynamics. 2. Equation of Continuity for Compressible Gas. 3. Equation of Gas Motion. 4. Equation of Energy. 5. Equation of Gasdynamics in Lagrange coordinates. 6. Boundary Conditions for Equations of Gasdynamics. 7. Some Peculiarities of Gasdynamics Models. Exercises.

CHAPTER III
MODELS DEDUCED FROM VARIATIONAL PRINCIPLES, HIERARCHIES OF MODELS

§ 1. Equations of Motion, Variational Principles and Conservation Laws in Classical Mechanics
1. Newtonian Equations of Motion for Mechanical System. 2. Lagrange Form of Equations of Motion. 3. Hamilton Variational Principle. 4. Conservation Laws and Space—Time Properties. Exercises.

§ 2. Models of Some Mechanical Systems
1. Pendulum on Free Suspension. 2. Nonpotential Oscillations. 3. Small Oscillations of a String. 4. Electro-Mechanical Analogy. Exercises.

§ 3. Boltzmann Equation and Derivative Equations
1. Description of Particles Ensemble by Distribution Function. 2. Boltzmann Equation for Distribution Function. 3. Maxwell Distribution and H-Theorem. 4. Equations for Distribution Function Moments. 5. A Chain of Hydrodynamic Models of Gas. Exercises.

CHAPTER IV
MODELS OF SOME HARD TO FORMALIZE OBJECTS

§ 1. Universality of Mathematical Models
1. Amoebas Gathering Dinamics. 2. Random Markov Process. 3. Examples of Analogies between Mechanical, Thermodynamic and Economic Processes. Exercises.

§ 2. Some Models of Financial and Economic Processes
1. Organizing of Publicity Campaign. 2. Mutual Debts Cancellation. 3. Macromodel of Market Economics Equilibrium. 4. Macromodel of Economic Growth. Exercises.

§ 3. Some Models of Competition
1. Interrelations in System "Pursuer—Victim". 2. Arms Race between Two Countries. 3. Combat Action of Rival Armies. Exercises.

§ 4. Power Distribution Dynamics in Hierarchy
1. General Statement of the Problem and Terminology. 2. Mechanisms of Power Redistribution in Hierarchical Structure. 3. Power Balance in Authority Institution, Conditions on Boundaries of Hierarchy, Passage to Continuos Model. 4. Law-Governed System "Power—Society". Stationary Distribution of Power and Violation of Authorized Limits. 5. Role of Main Properties of System in Legal Frameworks Violation. 6. Interpretation of Results and Conclusions. Exercises.

CHAPTER V
METHODS OF MATHEMATICAL MODELS INVESTIGATION

§ 1. Application of Similarity Methods
1. Dimensional and Group Analysis of Models. 2. Self-Similar Processes. 3. Various Regimes of Perturbations Propagation in Nonlinear Media. Exercises.

§ 2. Maximum Principle and Comparison Theorems
1. Formulation and Some Corollaries. 2. Blow-up Regimes Classification. 3. Extantion of Self-Simularity Method. Exercises.

§ 3. Theory of Averaging
1. Localized Structures in Nonlinear Media. 2. Various Ways of Averaging. 3. Classification of Combustion Regimes in Heat Conducting Media. Exercises.

§ 4. On Passage to Discrete Models
1. Inevitability of Numerical Modelling, Elementary Concepts of Difference Schemes Theory. 2. Direct Formal Approximation. 3. Integro-Interpolational Method. 4. Principle of Complete Conservativeness. 5. Difference Schemes Construction with the Help of Variational Principles. 6. Use of Hierarchical Approach to Discrete Models Building. Exercises.

CHAPTER VI
MATHEMATICAL MODELLING OF COMPLEX OBJECTS

§ 1. Problems of Technology and Ecology
1. Physically "Safe" Nuclear Reactor. 2. Hydrological Barrier Against Ground Waters Pollutions. 3. Complex Regimes of Gas Flow Around Bodies. 4. Biologically Acceptable Ways of Carbon—Hydrogen Fuels Burning.

§ 2. Fundamental Problems of Natural Sciences
1. Nonlinear Phenomena in Laser-Driven Thermonuclear Plasma. 2. Mathematical Restoration of Collision between Earth and Comet (Tunguska Event). 3. Climatic Consequences of Nuclear Conflict. 4. Magnetohydrodynamic Solar Dinamo.

§ 3. Computational Experiments with Models of Hard to Formalize Objects
1. Dissipative Biological Structures. 2. Processes in Economics of Transient Period. 3. Totalitarian and Anarchical Evolutions of Power Distribution in Hierarchies.

REFERENCES

The book includes 123 figures, 5 tables, 116 exercises, 89 references.